(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
5 June 2003 (05.06.2003)

PCT

(10) International Publication Number
WO 03/045990 A2

(51) International Patent Classification[7]: C07K 14/46, 14/47, C12N 1/16, 1/21, 5/10, 15/12, 15/63, C12Q 1/68, G01N 33/53

(21) International Application Number: PCT/EP02/13866

(22) International Filing Date:
26 November 2002 (26.11.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/333,348	26 November 2001 (26.11.2001)	US
60/384,537	31 May 2002 (31.05.2002)	US
60/422,471	30 October 2002 (30.10.2002)	US

(71) Applicant *(for all designated States except US)*: **HYBRIGENICS** [FR/FR]; 3/5 Impasse Reille, F-75014 Paris (FR).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **LEGRAIN, Pierre** [FR/FR]; 5, rue Mizon, F-75015 Paris (FR). **GAUTHIER, Jean-Michel** [FR/FR]; 9 Avenue du Maréchal Galliéni, F-78700 Conflans Sainte Honorine (FR). **COLLAND, Frédéric** [FR/FR]; 16, rue du Manoir, F-95380 Puiseux (FR). **JACQ, Xavier** [FR/FR]; 11 rue de Rambouillet, F-75012 Paris (FR).

(74) Agents: **ERNEST Gutmann-Yves Plasseraud S.A.** et al.; 3 rue Chauveau-Lagarde, F-75008 Paris (FR).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/045990 A2

(54) Title: PROTEIN-PROTEIN INTERACTIONS INVOLVING TRANSFORMING GROWTH FACTOR β SIGNALING OR INVOLVING TRANSDUCTION SIGNALS OF TRANSFORMING FACTOR β FAMILY MEMBERS

(57) Abstract: The present invention relates to protein-protein interactions involved in transforming growth factor β disorders and/or diseases. More specifically, the present invention relates to complexes of polypeptides or polynucleotides encoding the polypeptides, fragments of the polypeptides, antibodies to the complexes, Selected Interacting Domains (SID®) which are identified due to the protein-protein interactions, methods for screening drugs for agents which modulate the interaction of proteins and pharmaceutical compositions that are capable of modulating the protein-protein interactions.

PROTEIN-PROTEIN INTERACTIONS INVOLVING TRANSFORMING GROWTH FACTOR β SIGNALING OR INVOLVING TRANSDUCTION SIGNALS OF TRANSFORMING FACTOR β FAMILY MEMBERS

The present application claims priority to US provisional applications No. 60/333,348 filed on November 26, 2001, No. 60/384,537 filed on May 31, 2002 and No. 60/422,471 filed on October 30, 2002.

BACKGROUND AND PRIOR ART

Most biological processes involve specific protein-protein interactions. Protein-protein interactions enable two or more proteins to associate. A large number of non-covalent bonds form between the proteins when two protein surfaces are precisely matched. These bonds account for the specificity of recognition. Thus, protein-protein interactions are involved, for example, in the assembly of enzyme subunits, in antibody-antigen recognition, in the formation of biochemical complexes, in the correct folding of proteins, in the metabolism of proteins, in the transport of proteins, in the localization of proteins, in protein turnover, in first translation modifications, in the core structures of viruses and in signal transduction.

General methodologies to identify interacting proteins or to study these interactions have been developed. Among these methods are the two-hybrid system originally developed by Fields and co-workers and described, for example, in U.S. Patent Nos. 5,283,173, 5,468,614 and 5,667,973, which are hereby incorporated by reference.

The earliest and simplest two-hybrid system, which acted as basis for development of other versions, is an *in vivo* assay between two specifically constructed proteins. The first protein, known in the art as the "bait protein" is a chimeric protein which binds to a site on DNA upstream of a reporter gene by means of a DNA-binding domain or BD. Commonly, the binding domain is the DNA-binding domain from either Gal4 or native *E. coli* LexA and the sites placed upstream of the reporter are Gal4 binding sites or LexA operators, respectively.

The second protein is also a chimeric protein known as the "prey" in the art. This second chimeric protein carries an activation domain or AD. This activation domain is typically derived from Gal4, from VP16 or from B42.

Besides the two-hybrid systems, other improved systems have been developed to detected protein-protein interactions. For example, a two-hybrid plus one system was developed that allows the use of two proteins as bait to screen available cDNA libraries to detect a third partner. This method permits the detection between proteins that are part of a larger protein complex such as the RNA polymerase II holoenzyme and the TFIIH or TFIID complexes. Therefore, this method, in general, permits the detection of ternary complex

formation as well as inhibitors preventing the interaction between the two previously defined fused proteins.

Another advantage of the two-hybrid plus one system is that it allows or prevents the formation of the transcriptional activator since the third partner can be expressed from a conditional promoter such as the methionine-repressed Met25 promoter which is positively regulated in medium lacking methionine. The presence of the methionine-regulated promoter provides an excellent control to evaluate the activation or inhibition properties of the third partner due to its "on" and "off" switch for the formation of the transcriptional activator. The three-hybrid method is described, for example in Tirode *et al.*, *The Journal of Biological Chemistry*, **272**, No. 37 pp. 22995-22999 (1997) incorporated herein by reference.

Besides the two and two-hybrid plus one systems, yet another variant is that described in Vidal et al, *Proc. Natl. Sci. 93* pgs. 10315-10320 called the reverse two- and one-hybrid systems where a collection of molecules can be screened that inhibit a specific protein-protein or protein-DNA interactions, respectively.

A summary of the available methodologies for detecting protein-protein interactions is described in Vidal and Legrain, *Nucleic Acids Research* Vol. 27, No. 4 pgs. 919-929 (1999) and Legrain and Selig, FEBS Letters 480 pgs. 32-36 (2000) which references are incorporated herein by reference.

However, the above conventionally used approaches and especially the commonly used two-hybrid methods have their drawbacks. For example, it is known in the art that, more often than not, false positives and false negatives exist in the screening method. In fact, a doctrine has been developed in this field for interpreting the results and in common practice an additional technique such as co-immunoprecipitation or gradient sedimentation of the putative interactors from the appropriate cell or tissue type are generally performed. The methods used for interpreting the results are described by Brent and Finley, Jr. in *Ann. Rev. Genet., 31 pgs. 663-704 (1997).* Thus, the data interpretation is very questionable using the conventional systems.

One method to overcome the difficulties encountered with the methods in the prior art is described in WO99/42612, incorporated herein by reference. This method is similar to the two-hybrid system described in the prior art in that it also uses bait and prey polypeptides. However, the difference with this method is that a step of mating at least one first haploid recombinant yeast cell containing the prey polypeptide to be assayed with a second haploid recombinant yeast cell containing the bait polynucleotide is performed. Of course the person skilled in the art would appreciate that either the first recombinant yeast cell or the second recombinant yeast cell also contains at least one detectable reporter gene that is activated by a polypeptide including a transcriptional activation domain.

The method described in WO99/42612 permits the screening of more prey polynucleotides with a given bait polynucleotide in a single step than in the prior art systems due to the cell to cell mating strategy between haploid yeast cells. Furthermore, this method is more thorough and reproducible, as well as sensitive. Thus, the presence of false negatives and/or false positives is extremely minimal as compared to the conventional prior art methods.

Transforming growth factor β (TGFβ) belongs to a super-family of cytokines, including TGFβ1, TGFβ2, TGFβ3, activins and Bone Morphologenetic Proteins (hereinafter BMP), which are synthesized by many cell types and have a variety of cellular and biological effects, including control of proliferation, differentiation, migration, angiogenesis, immunity and regulation of the turnover of the extracellular matrix. A number of disease states are known to be associated with variations in expression of genes which are controlled by TGFβ and related ,cytokines, including fibrotic disorders, abnormal wound healing, abnormal bone formation, cancer and tumor development, neurologic disorders, haematopoiesis and immune and inflammatory disorders.

Signaling by this family of cytokines is transduced by heteromeric complexes of transmembrane Ser/Thr kinase receptors. Upon ligand binding, type II receptor phosphorylates and activates type I receptor which then propagates signals to downstream targets, in particular the Smad proteins.

Ten mammalian Smad proteins have been identified and divided into three classes. The first includes pathway-restricted proteins such as Smad1, Smad5 and Smad8 which are specifically involved in BMP signaling and Smad2 and Smad3 which are restricted to TGFβ/activin pathway. The second class contains the common-mediator Smad4 implicated in both BMP and TGFβ/activin pathways. The third class contains the inhibitory Smads, Smad6 and Smad7. At least Smad2 and Smad3 are retained in the cytoplasm by binding to the SARA protein. After phosphorylation by TGFβ-activated type I receptor on their carboxy-terminal SSXS sequence, pathway-restricted Smads form heteromeric complexes with Smad4 and then translocate to the nucleus where they control expression of diverse genes involved in various biological processes such as control of cellular proliferation and differentiation, regulation of the immune system and regulation of the extracellular matrix formation.

Several proteins such as TGIF, Ski, SnoN, SNIP1 and CBP have been identified as Smad transcriptional co-regulators and shown to modulate the transcriptional ability of Smad proteins by direct interactions. Finally, proteins such Smurf1 and Smurf2 are involved in degradation of Smad proteins by the proteasome machinery.

Most biological processes involve specific protein-protein interactions. Protein-protein interactions enable two or more proteins to associate. A large number of non-covalent bonds

form between the proteins when two protein surfaces are precisely matched. These bonds account for the specificity of recognition. Thus, protein-protein interactions are involved, for example, in the assembly of enzyme subunits, in antibody-antigen recognition, in the formation of biochemical complexes, in the correct folding of proteins, in the metabolism of proteins, in the transport of proteins, in the localization of proteins, in protein turnover, in first translation modifications, in the core structures of viruses and in signal transduction.

Several members of the TGFβ/BMP pathways (SARA, Smurf1, Smurf2, Smad1, Smad2/hMAD2, Smad3/hMAD-3, Smad4, Smad5/MADH5, Smad7, Smad9/MADH6, SNIP1, SnoN) have been used as baits in yeast-two hybrid screening experiments. Several proteins have been identified as interactors with thoses baits (Figure 10). It was showed here functional data in mammalian cells that validate that those interactants are proteins involved in TGFβ/BMP signaling.

Thus, there is the still a need to explore all mechanisms relating to transforming growth factor β protein and to identify drug targets for fibrotic disorders, abnormal wound healing, abnormal bone formation, cancer and tumor development, neurologic disorders, haematopoiesis and immune and inflammatory disorders and/or diseases.

SUMMARY OF THE PRESENT INVENTION

Thus, it is an aspect of the present invention to identify protein-protein interactions involving proteins of the transforming growth factor β super-family of cytokines transduction pathway and to identify drug targets for fibrotic disorders, abnormal wound healing, abnormal bone formation, cancer and tumor development, neurologic disorders, haematopoiesis and immune and inflammatory disorders and/or disease.

It is another aspect of the present invention to identify protein-protein interactions involved in transforming growth factor β-mediated disorders and/or diseases for the development of more effective and better targeted therapeutic treatments.

It is yet another aspect of the present invention to identify complexes of polypeptides or polynucleotides encoding the polypeptides and fragments of the polypeptides of the transforming growth factor β super-family of cytokines transduction pathway.

It is yet another aspect of the present invention to identify antibodies to these complexes of polypeptides or polynucleotides encoding the polypeptides and fragments of the polypeptides involving transforming growth factor β signaling including polyclonal, as well as monoclonal antibodies that are used for detection.

It is still another aspect of the present invention to identify selected interacting domains of the polypeptides, called SID® polypeptides.

It is still another aspect of the present invention to identify selected interacting domains of the polynucleotides, called SID® polynucleotides.

It is still another aspect of the present invention to provide a diagnostic kit to test for deficiencies in the transforming growth factor β super-family of cytokines transduction pathway.

It is another aspect of the present invention to identify interacting proteins in the transforming growth factor β super-family of cytokines transduction pathway that can be used in pharmaceutical compositions or for diagnostic purposes.

It is another aspect of the present invention to generate protein-protein interactions maps called PIM®s.

It is yet another aspect of the present invention to provide a method for screening drugs for agents which modulate the interaction of proteins and pharmaceutical compositions that are capable of modulating the protein-protein interactions involved in transforming growth factor β disorders and/or diseases.

It is another aspect to administer the nucleic acids of the present invention via gene therapy.

It is yet another aspect of the present invention to provide protein chips or protein microarrays.

It is yet another aspect of the present invention to provide a report in, for example paper, electronic and/or digital forms, concerning the protein-protein interactions, the modulating compounds and the like as well as a PIM®.

These and other aspects are achieved by the present invention as evidenced by the summary of the invention, description of the preferred embodiments and the claims.

Thus the present invention relates to a complex of interacting proteins of columns 1 and 4 of Table 2.

Furthermore, the present invention provides SID® polynucleotides and SID® polypeptides of Table 3, as well as a PIM® involved in transforming growth factor β-mediated disorders and/or diseases.

The present invention also provides antibodies to the protein-protein complexes involved in transforming growth factor β-mediated disorders and/or diseases.

In another embodiment the present invention provides a method for screening drugs for agents that modulate the protein-protein interactions and pharmaceutical compositions that are capable of modulating protein-protein interactions.

In another embodiment the present invention provides protein chips or protein microarrays.

In yet another embodiment the present invention provides a report in, for example, paper, electronic and/or digital forms.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a schematic representation of the pB6 plasmid.

Fig. 2 is a schematic representation of the pB20 plasmid.

Fig. 3 is a schematic representation of the pP6 plasmid.

Fig. 4 is a schematic representation of vectors expressing the T25 fragment.

Fig. 5 is a schematic representation of vectors expressing the T18 fragment.

Fig. 6 is a schematic representation of various vectors of pCmAHL1, pT25 and pT18.

Fig. 7 is a schematic representation identifying the SID®'s of proteins of the present invention. In this figure the "Full-length prey protein" is the Open Reading Frame (ORF) or coding sequence (CDS) where the identified prey polypeptides are included. The Selected Interaction Domain (SID®) is determined by the commonly shared polypeptide domain of every selected prey fragment.

Fig. 8 is a protein map (PIM®).

Fig. 9 is a schematic representation of the pB27 plasmid.

Fig. 10 is a schematic representation of the pB28 plasmid.

Fig. 11 is a schematic representation of a protein interaction map around the newly functionally characterized proteins described in the present invention. These 10 proteins are highlighted by the symbol "*". The Predicted Biological Score (PBS) is represented by a code on each line and classified from A to E (Rain *et al.*, 2001). PP1ca is also named PPP1CA. MADH5 and MADH6 correspond to Smad5 and Smad9, respectively. hMAD-2 and h-MAD-3 correspond to Smad2 and Smad3, respectively. MAN1 is the orthologous of SANE, a protein recently identified as involved in the BMP pathway (Raju *et al.*, 2002)

Fig. 12 is a schematic representation of a protein interaction map between ZNF8 and Smad proteins. The full-length proteins are represented in grey and black boxes correspond to the interaction domains. Using two-hybrid screening, ZNF8 was shown to interact with Smad1 (A), Smad4 (B), Smad5 (C) and Smad9 (D). Amino-acid position are indicated.

Fig. 13 A, B and C are graphs showing that ZNF8 siRNA represses TGFβ- and BMP-dependent luciferase reporter activities. HepG2 cells were transiently transfected in 24 well-plates as described under *Materials & Methods* with the BMP reponsive luciferase reporter, $p(GC)_{12}$-MLP-Luc (A & B) or the TGFβ responsive luciferase reporter, $p(GTCT)_8$-MLP-Luc (C). All experiments included pRL-TK as an internal transfection control. A TβRI-targeting siRNA duplex was used as a positive control for disruption of the TGFβ pathway. A mutated version of the TβRI-targeting siRNA duplex (2 mismatches versus consensus sequence) was used as a negative control. SiRNA transfections were performed at 4 and 40nM. Co-transfection of ZNF8-targeting siRNA duplex was tested in cells treated or not with 50ng/ml BMP7 (A), 50ng/ml BMP6 (B) or 5 ng/ml TGFβ1 (C) for 18 hours in cells pre-starved for 2 hours in serum-free culture medium. Cells were harvested 48 hours after transfection and 10µl of lysates were used for the *Dual Luciferase Assay*. Data are representative of two or

three independant duplicated experiments and are presented as a ratio between firefly and renilla luciferases.

Fig. 14A, B and C are graphs showing that ZNF8 siRNA specifically represses BMP-dependent markers. HepG2 cells were transiently transfected in 24 well-plates as described under *Materials & Methods* with a control siRNA (pGL3-targeting siRNA) or ZNF8-targeting siRNA duplex. Cells were treated or not with 50ng/ml of recombinant human BMP7 for 18 hours in cells pre-starved for 2 hours in serum-free culture medium. SiRNA transfections were performed either at 0.5nM and 2.5nM (A & B) or at 4 and 40nM (C) of duplex. Cells were harvested and lysed 48 hours after transfection. Total RNA were extracted as described under *Materials &* Methods and quantitative PCR analysis were performed in order to quantitate the endogenous levels of the BMP pathway markers junB (A) and alkaline phosphatase (B& C). Data are representative of two or three independant duplicated experiments and are presented as normalized RNA levels using either GAPDH (A & B) or hGUS (C).

Fig. 15 A and B are graphs showing that ZNF8 siRNA does not repress BMP-independent markers. HepG2 cells were transiently transfected in 24 well-plates as described under *Materials & Methods* with a control siRNA (pGL3-targeting siRNA) or ZNF8-targeting siRNA duplex. Cells were treated or not with 50ng/ml of recombinant human BMP7 for 18 hours in cells pre-starved for 2 hours in serum-free culture medium. SiRNA transfections were performed either at 0.5nM and 2.5nM (A) or at 4 and 40nM (B) of duplex. Cells were harvested and lysed 48 hours after transfection. Total RNA were extracted as described under *Materials &* Methods and quantitative PCR analysis were performed in order to quantitate the endogenous levels of the TGFβ pathway marker PAI-1 (PAI-1 hereinafter Plasminogen Activator Inhibitor I) (A) and an unrelated marker, hGUS (B). Data are representative of two or three independant duplicated experiments and are presented as normalized RNA levels using either GAPDH (A) or relative levels (B).

Fig. 16 is a schematic representation of an Interaction between LAPTm5 and Smurf2. The full-length proteins are represented in grey and black boxes correspond to the interaction domains. Using two-hybrid screening, interaction between Smurf2 and LAPTm5 was found in both directions. Smurf2 was shown to interact with the C-terminal domain of LAPTm5.

Fig. 17 A and B are graphs showing that LAPTm5 specifically inhibits the TGFβ pathway. The effect of LAPTm5 over-expression was studied using the following Luciferase reporter vectors: a TGFβ responsive element (TGF-RE = $p(GTCT)_6$-MLP-Luc), a BMP-responsive element (BMP-RE = $p(GC)_{12}$-MLP-Luc) and an unrelated reporter (pGL3 control) (see Materials & Methods). The effect was studied in the presence or absence of TGFβ (10 ng/ml) or BMP7 (50 ng/ml), as described. This study was performed with 0, 2 or 10 ng of

pV3-LAPTm5 in HepG2 cells (A) or with 0, 0.5, 2, 10 or 50 ng of pV3-LAPTm5 in HEK293 cells (B). The specific Luciferase activity was normalized using the pRL-TK vector. Experiments were performed in triplicate.

Fig. 18 A and B are graphs showing that LAPTm5 expression is up-regulated by TGFβ The endogenous level of LAPTm5 mRNA was determined in several cell lines by Q-PCR experiments using the LAPTm5 probe (see Materials & Methods). Ct levels of LAPTm5 mRNA is given for each cell lines (A). The endogenous level of mRNA was determined in HepG2 cells in the presence or absence of TGFβ (10 ng/ml) with or without a TβRI-targeting siRNA duplex (B) (TβRI hereinafter Transforming Growth Factor β Receptor I.

Fig. 19 A and B are graphs showing that LAPTm5 siRNA up-regulates BMP and TGFβ-dependent reporter activities. HepG2 cells were transiently transfected in 24 well-plates as described under *Materials & Methods* with the TGFβ reponsive luciferase reporter, $p(GTCT)_8$-MLP-Luc (A) or the BMP responsive luciferase reporter, $p(GC)_{12}$-MLP-Luc (B). All experiments included pRL-TK as an internal transfection control. A TβRI-targeting siRNA duplex was used as a positive control for disruption of the TGFβ pathway. A mutated version of the TβRI-targeting siRNA duplex (2 mismatches versus consensus sequence) was used as a negative control. SiRNA transfections were performed at 4 and 40nM. Co-transfection of LAPTm5-targeting siRNA duplex was tested in cells treated or not with 5ng/ml recombinant human TGFβ (A), 50ng/ml recombinant human BMP7 (B) for 18 hours in cells pre-starved for 2 hours in serum-free culture medium. Cells were harvested 48 hours after transfection and 10µl of lysates were used for the *Dual Luciferase Assay*. Data are representative of two or three independant duplicated experiments and are presented as a ratio between firefly and renilla luciferases.

Fig. 20 A, B, C and D are graphs showing that LAPTm5 siRNA up-regulates BMP and TGFβ-dependent markers. HepG2 cells were transiently transfected in 24 well-plates as described under *Materials & Methods* with a control siRNA (pGL3-targeting siRNA) or LAPTm5-targeting siRNA duplex. Cells were treated or not with 5 ng/ml of recombinant human TGFβ1 or 50ng/ml of recombinant human BMP7 for 18 hours in cells pre-starved for 2 hours in serum-free culture medium. SiRNA transfections were performed at 40nM of duplex (A, B, C & D). Cells were harvested and lysed 48 hours after transfection. Total RNA were extracted as described under *Materials &* Methods and quantitative PCR analysis were performed in order to quantitate the endogenous levels of the TGFβ pathway markers PAI-1 and junB (A & B, respectively) and a BMP pathway marker, alkaline phosphatase (C). Data are representative of two or three independant duplicated experiments and are presented as normalized RNA levels using hGUS (A, B & C). Relative levels of hGUS in the same experiment are also shown (D).

Fig. 21 is a schematic representation of an Interaction between RNF11 Smurf1, Smurf2 and SARA.The full-length proteins are represented in grey and black boxes correspond to the interaction domains. Using two-hybrid screening, RNF11 was shown to interact with Smurf1 (A), Smurf2 (B), and SARA (C). Amino-acid positions are indicated.

Fig. 22 is a gel showing that RNF11 is involved in regulating SARA protein levels. Transfection experiments with pV3-SARA (200 ng) and/or pV3-RNF11 (300 ng) in the presence or absence of TGFβ (10 ng/ml) were performed. After TGFβ induction for 18H, cells' lysates were resolved on a 4-12% NuPAGE gradient gel, transfered and revealed using anti-SARA antibody (see Materials & Methods).

Fig. 23 is a schematic diagram showing the Interaction between KIAA1196 and Smad1. The full-length proteins are represented in grey and black boxes correspond to the interaction domains. Using two-hybrid screening, KIAA1196 was shown to interact with Smad1.

Fig. 24 A and B are graphs showing that KIAA1196 siRNA specifically represses TGFβ-dependent markers in HepG2 cells.HepG2 cells were transiently transfected in 24 well-plates as described under *Materials & Methods* with the TGFβ responsive luciferase reporter, $p(GTCT)_8$-MLP-Luc (A) or the BMP reponsive luciferase reporter, $p(GC)_{12}$-MLP-Luc (B). All experiments included pRL-TK as an internal transfection control. A TβRI-targeting siRNA duplex was used as a positive control for disruption of the TGFβ pathway. A mutated version of the T RI-targeting siRNA duplex (2 mismatches versus consensus sequence) was used as a negative control. SiRNA transfections were performed at 4 and 40nM. Co-transfection of KIAA1196-targeting siRNA duplex was tested in cells treated or not with 5ng/ml recombinant human TGFβ (A) and 50ng/ml recombinant human BMP6 (B) for 18 hours in cells pre-starved for 2 hours in serum-free culture medium. Cells were harvested 48 hours after transfection and 10µl of lysates were used for the *Dual Luciferase Assay*. Data are representative of two or three independant duplicated experiments and are presented as a ratio between firefly and renilla luciferases.

Fig. 25 is a graph showing that KIAA1196 siRNA specifically represses TGFβ-dependent reporter activity in HEK293 cells. HEK 293 cells were transiently transfected in 24 well-plates as described under *Materials & Methods* with the TGFβ responsive luciferase reporter, $p(GTCT)_8$-MLP-Luc. All experiments included pRL-TK as an internal transfection control. A TβRI-targeting siRNA duplex was used as a positive control for disruption of the TGFβ pathway. A mutated version of the TβRI-targeting siRNA duplex (2 mismatches versus consensus sequence) was used as a negative control. SiRNA transfections were performed at 30nM. Co-transfection of KIAA1196-targeting siRNA duplex was tested in cells treated or not with 5ng/ml recombinant human TGFβ for 18 hours in cells pre-starved for 2 hours in serum-free culture medium. Cells were harvested 48 hours after transfection and 10µl of

lysates were used for the *Dual Luciferase Assay*. Data are representative of two or three independant duplicated experiments and are presented as a ratio between firefly and renilla luciferases.

Fig. 26 A, B, C and D are graphs showing that KIAA1196 siRNA specifically represses TGFβ-dependent markers. HepG2 cells were transiently transfected in 24 well-plates as described under *Materials & Methods* with a control siRNA (pGL3-targeting siRNA) or KIAA1196-targeting siRNA duplex. Cells were treated or not with 5 ng/ml of recombinant human TGFβ1 or 50ng/ml of recombinant human BMP7 for 18 hours in cells pre-starved for 2 hours in serum-free culture medium. SiRNA transfections were performed at 40nM of duplex (A, B, C & D). Cells were harvested and lysed 48 hours after transfection. Total RNA were extracted as described under *Materials &* Methods and quantitative PCR analysis were performed in order to quantitate the endogenous levels of the TGFβ pathway markers PAI-1 and junB (A & B, respectively) and a BMP pathway marker, alkaline phosphatase (C). Data are representative of two or three independant duplicated experiments and are presented as normalized RNA levels using hGUS (A, B & C). Relative levels of hGUS in the same experiment are also shown (D).

Fig. 27 is a schematic representation showing the Interaction between LMO4 and Smad9. The full-length proteins are represented in grey and black boxes correspond to the interaction domains. Using two-hybrid screening, LMO4 was shown to interact with Smad9.

Fig. 28 A, B and C are graphs showing that LMO4 siRNA specifically repress a BMP-dependent luciferase reporter. HepG2 cells were transiently transfected in 24 well-plates as described under *Materials & Methods* with the BMP reponsive luciferase reporter, $p(GC)_{12}$-MLP-Luc (A) or the TGFβ responsive luciferase reporter, $p(GTCT)_8$-MLP-Luc (B). All experiments included pRL-TK as an internal transfection control. A TβRI-targeting siRNA duplex was used as a positive control for disruption of the TGF pathway. A mutated version of the TβRI-targeting siRNA duplex (2 mismatches versus consensus sequence) was used as a negative control. SiRNA transfections were performed at 4 and 40nM. Co-transfection of LMO4-targeting siRNA duplex was tested in cells treated or not with 50ng/ml recombinant human BMP7 or BMP6 (A & B, respectively) and 5ng/ml recombinant human TGFβ (C) for 18 hours in cells pre-starved for 2 hours in serum-free culture medium. Cells were harvested 48 hours after transfection and 10µl of lysates were used for the *Dual Luciferase Assay*. Data are. representative of two or three independant duplicated experiments and are presented as a ratio between firefly and renilla luciferases.

Fig. 29 A and B are graphs showing that LMO4 siRNA specifically represses BMP-induced markers in BMP7-treated HepG2 cells.HepG2 cells were transiently transfected in 24 well-plates as described under *Materials & Methods* with a control siRNA (pGL3-targeting siRNA) or LMO4-targeting siRNA duplex. Cells were treated or not with 50ng/ml of

recombinant human BMP7 for 18 hours in cells pre-starved for 2 hours in serum-free culture medium. SiRNA transfections were performed at 0.5 or 2.5nM of duplex (A) and 4 or 40nM of duplex (B). Cells were harvested and lysed 48 hours after transfection. Total RNA were extracted as described under *Materials* & Methods and quantitative PCR analysis were performed in order to quantitate the endogenous levels of the BMP pathway marker alkaline phosphatase (A & B). Data are representative of two or three independant duplicated experiments and are presented as normalized RNA levels using hGUS (A, B).

Fig. 30 A, B and C are graphs showing that LMO4 siRNA does not repress BMP-independent markers in BMP7-treated HepG2 cells.HepG2 cells were transiently transfected in 24 well-plates as described under *Materials & Methods* with a control siRNA (pGL3-targeting siRNA) or LMO4-targeting siRNA duplex. Cells were treated or not with 50ng/ml of recombinant human BMP7 for 18 hours in cells pre-starved for 2 hours in serum-free culture medium. SiRNA transfections were performed at 4 or 40nM of duplex (A, B) and 0.5 or 2.5nMof duplex (C). Cells were harvested and lysed 48 hours after transfection. Total RNA were extracted as described under *Materials* & Methods and quantitative PCR analysis were performed in order to quantitate the endogenous levels of the TGFβ and BMP pathways marker junB (A) and a TGFβ pathway marker, PAI-1 (C). Data are representative of two or three independant duplicated experiments and are presented as normalized RNA levels using hGUS (A) or using GAPDH (C). Relative levels of hGUS in the same experiment are also shown (B).

Fig. 31 is a schematic diagram showing the interaction between PP1ca and SARA. The full-length proteins are represented in grey and black boxes correspond to the interaction domains. Using two-hybrid screening, PP1ca was shown to interact with SARA.

Fig. 32 A and B are graphs showing that PP1ca stimulates the TGFβ pathway. The effect of PP1ca over-expression was studied using the following luciferase reporter vectors: a TGFβ responsive element (TGF-RE = $p(GTCT)_8$-MLP-Luc), a BMP-responsive element (BMP-RE = $p(GC)_{12}$-MLP-Luc) and an unrelated reporter (pGL3 control) (see Materials & Methods). The effect was studied in the presence or absence of TGFβ (10 ng/ml) or BMP7 (50 ng/ml), as described. This study was performed with 0, 10, 50 or 200 ng of pV3-PP1ca in HepG2 cells (A) or in HEK293 cells (B). The specific Luciferase activity was normalized using the pRL-TK vector. Experiments were performed in triplicate.

Fig. 33 A, B and C are graphs showing that PP1ca stimulates PAI-1 mRNA expression. Baculoviruses containing the Smad3 or PP1ca genes under the control of the CMV promoter were generated and used to infect HepG2 cells (see Materials & Methods). The over-expression level was checked and quantified by Q-PCR (A). The endogenous PAI-1 mRNA levels were measured by Q-PCR 24 hours post infection with Smad3 or PP1ca-containing

baculoviruses in the presence or absence of TGFβ (10 ng/ml). The value 1 is attributed to the mRNA amount of PAI-1 in the absence of TGFβ and in the absence of infection (B).

Fig. 34 is a schematic diagram showing the Interaction between HYPA and Smad4. The full-length proteins are represented in grey and black boxes correspond to the interaction domains. Using two-hybrid screening, HYPA was shown to interact with Smad4.

Fig. 35 A, B and C are graphs showing that HYPA siRNA specifically represses BMP-dependent reporter activity.HepG2 cells were transiently transfected in 24 well-plates as described under *Materials & Methods* with the BMP reponsive luciferase reporter, $p(GC)_{12}$-MLP-Luc (A & B) or the TGFβ responsive luciferase reporter, $p(GTCT)_8$-MLP-Luc (C). All experiments included pRL-TK as an internal transfection control. A TβRI-targeting siRNA duplex was used as a positive control for disruption of the TGFβ pathway. A mutated version of the TβRI-targeting siRNA duplex (2 mismatches versus consensus sequence) was used as a negative control. SiRNA transfections were performed at 4 and 40nM. Co-transfection of HYPA-targeting siRNA duplex was tested in cells treated or not with 50ng/ml recombinant human BMP7 or BMP6 (A & B, respectively) and 5ng/ml recombinant human TGFβ (C) for 18 hours in cells pre-starved for 2 hours in serum-free culture medium. Cells were harvested 48 hours after transfection and 10µl of lysates were used for the *Dual Luciferase Assay*. Data are representative of two or three independant duplicated experiments and are presented as a ratio between firefly and renilla luciferases.

Fig. 36 is a graph showing that HYPA siRNA represses BMP-dependent markers. HepG2 cells were transiently transfected in 24 well-plates as described under *Materials & Methods* with a control siRNA (pGL3-targeting siRNA) or HYPA-targeting siRNA duplex. Cells were treated or not with 50ng/ml of recombinant human BMP7 for 18 hours in cells pre-starved for 2 hours in serum-free culture medium. SiRNA transfections were performed at 0.5 or 2.5nM of duplex. Cells were harvested and lysed 48 hours after transfection. Total RNA were extracted as described under *Materials &* Methods and quantitative PCR analysis were performed in order to quantitate the endogenous levels of the BMP pathway marker alkaline phosphatase. Data are representative of two or three independant duplicated experiments and are presented as normalized RNA levels using GAPDH.

Fig. 37 is a schematic diagram showing the Interaction between FLJ20037 and SARA. The full-length proteins are represented in grey and black boxes correspond to the interaction domains. Using two-hybrid screening, FLJ20037 was shown to interact with SARA.

Fig. 38 A, B and C are graphs showing that FLJ20037 stimulates PAI-1 mRNA expression. Baculoviruses containing the Smad3 or FLJ20037 genes under the control of the CMV promoter were generated and used to infect HepG2 cells (see Materials & Methods). The over-expression level was checked and quantified by Q-PCR (A). The endogenous PAI-

1 mRNA levels were measured by Q-PCR 24 hours post infection with Smad3 or FLJ20037-containing baculoviruses in the presence or absence of TGFβ (10 ng/mL). The value 1 is attributed to the mRNA amount of PAI-1 in the absence of TGFβ and in the absence of infection (B).

Fig. 39 is a graph showing that FLJ20037 siRNA down-regulates TGFβ-dependent markers.HepG2 cells were transiently transfected in 24 well-plates as described under *Materials & Methods* with a control siRNA (pGL3-targeting siRNA) or FLJ20037-targeting siRNA duplex. Cells were treated or not with 5ng/ml of recombinant human TGFβ for 18 hours in cells pre-starved for 2 hours in serum-free culture medium. SiRNA transfections were performed at 0.5 or 2.5nM of duplex. Cells were harvested and lysed 48 hours after transfection. Total RNA was extracted as described under *Materials &* Methods and quantitative PCR analysis were performed in order to quantitate the endogenous levels of the TGFβ pathway marker PAI-1. Data are representative of two or three independant duplicated experiments and are presented as normalized RNA levels using GAPDH.

Fig. 40 is a schematic diagram showing the Interaction between PTPN12 and Smad5. The full-length proteins are represented in grey and black boxes correspond to the interaction domains. Using two-hybrid screening, PTPN12 was shown to interact with Smad5. Amino-acid positions are indicated.

Fig. 41 A and B are graphs showing that PTPN12 siRNA up-regulates BMP and TGFβ-dependent reporter activities. HepG2 cells were transiently transfected in 24 well-plates as described under *Materials & Methods* with the BMP reponsive luciferase reporter, $p(GC)_{12}$-MLP-Luc (A) or the TGFβ responsive luciferase reporter, $p(GTCT)_8$-MLP-Luc (B). All experiments included pRL-TK as an internal transfection control. A TβRI-targeting siRNA duplex was used as a positive control for disruption of the TGF pathway. A mutated version of the TβRI-targeting siRNA duplex (2 mismatches versus consensus sequence) was used as a negative control. SiRNA transfections were performed at 4 and 40nM. Co-transfection of PTPN12-targeting siRNA duplex was tested in cells treated or not with 50ng/ml recombinant human BMP6 (A) and 5ng/ml recombinant human TGFβ (B) for 18 hours in cells pre-starved for 2 hours in serum-free culture medium. Cells were harvested 48 hours after transfection and 10µl of lysates were used for the *Dual Luciferase Assay*. Data are representative of two or three independant duplicated experiments and are presented as a ratio between firefly and renilla luciferases.

Fig. 42 A and B are schematic diagrams showing the Interaction between HIPK3, SnoN and SNIP1. The full-length proteins are represented in grey and black boxes correspond to the interaction domains. Using two-hybrid screening, HIPK3 was shown to interact with the N-terminal domains of SNIP1 (A) and SnoN (B). Amino-acid positions are indicated.

Fig. 43 A and B are graphs showing that HIPK3 siRNA specifically up-regulates BMP-dependent reporter activities.

HepG2 cells were transiently transfected in 24 well-plates as described under *Materials & Methods* with the BMP reponsive luciferase reporter, $p(GC)_{12}$-MLP-Luc (A) or the TGFβ responsive luciferase reporter, $p(GTCT)_8$-MLP-Luc (B). All experiments included pRL-TK as an internal transfection control. A T RI-targeting siRNA duplex was used as a positive control for disruption of the TGF pathway. A mutated version of the TβRI-targeting siRNA duplex (2 mismatches versus consensus sequence) was used as a negative control. SiRNA transfections were performed at 4 and 40nM. Co-transfection of HIPK3-targeting siRNA duplex was tested in cells treated or not with 50ng/ml recombinant human BMP6 (A) and 5ng/ml recombinant human TGFβ (B) for 18 hours in cells pre-starved for 2 hours in serum-free culture medium. Cells were harvested 48 hours after transfection and 10µl of lysates were used for the *Dual Luciferase Assay*. Data are representative of two or three independant duplicated experiments and are presented as a ratio between firefly and renilla luciferases.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

As used herein the terms "polynucleotides", "nucleic acids" and "oligonucleotides" are used interchangeably and include, but are not limited to RNA, DNA, RNA/DNA sequences of more than one nucleotide in either single chain or duplex form. The polynucleotide sequences of the present invention may be prepared from any known method including, but not limited to, any synthetic method, any recombinant method, any *ex vivo* generation method and the like, as well as combinations thereof.

The term "polypeptide" means herein a polymer of amino acids having no specific length. Thus, peptides, oligopeptides and proteins are included in the definition of "polypeptide" and these terms are used interchangeably throughout the specification, as well as in the claims. The term "polypeptide" does not exclude post-translational modifications such as polypeptides having covalent attachment of glycosyl groups, acetyl groups, phosphate groups, lipid groups and the like. Also encompassed by this definition of "polypeptide" are homologs thereof.

By the term "homologs" is meant structurally similar genes contained within a given species, orthologs are functionally equivalent genes from a given species or strain, as determined for example, in a standard complementation assay. Thus, a polypeptide of interest can be used not only as a model for identifying similiar genes in given strains, but also to identify homologs and orthologs of the polypeptide of interest in other species. The orthologs, for example, can also be identified in a conventional complementation assay. In addition or alternatively, such orthologs can be expected to exist in bacteria (or other kind of

cells) in the same branch of the phylogenic tree, as set forth, for example, at ftp://ftp.cme.msu.edu/pub/rdp/SSU-rRNA/SSU/Prok.phylo.

As used herein the term "prey polynucleotide" means a chimeric polynucleotide encoding a polypeptide comprising (i) a specific domain; and (ii) a polypeptide that is to be tested for interaction with a bait polypeptide. The specific domain is preferably a transcriptional activating domain.

As used herein, a "bait polynucleotide" is a chimeric polynucleotide encoding a chimeric polypeptide comprising (i) a complementary domain; and (ii) a polypeptide that is to be tested for interaction with at least one prey polypeptide. The complementary domain is preferably a DNA-binding domain that recognizes a binding site that is further detected and is contained in the host organism.

As used herein "complementary domain" is meant a functional constitution of the activity when bait and prey are interacting; for example, enzymatic activity.

As used herein "specific domain" is meant a functional interacting activation domain that may work through different mechanisms by interacting directly or indirectly through intermediary proteins with RNA polymerase II or III-associated proteins in the vicinity of the transcription start site.

As used herein the term "complementary" means that, for example, each base of a first polynucleotide is paired with the complementary base of a second polynucleotide whose orientation is reversed. The complementary bases are A and T (or A and U) or C and G.

The term "sequence identity" refers to the identity between two peptides or between two nucleic acids. Identity between sequences can be determined by comparing a position in each of the sequences which may be aligned for the purposes of comparison. When a position in the compared sequences is occupied by the same base or amino acid, then the sequences are identical at that position. A degree of sequence identity between nucleic acid sequences is a function of the number of identical nucleotides at positions shared by these sequences. A degree of identity between amino acid sequences is a function of the number of identical amino acid sequences that are shared between these sequences. Since two polypeptides may each (i) comprise a sequence (i.e., a portion of a complete polynucleotide sequence) that is similar between two polynucleotides, and (ii) may further comprise a sequence that is divergent between two polynucleotides, sequence identity comparisons between two or more polynucleotides over a "comparison window" refers to the conceptual segment of at least 20 contiguous nucleotide positions wherein a polynucleotide sequence may be compared to a reference nucleotide sequence of at least 20 contiguous nucleotides and wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) of 20 percent or less compared to the reference

sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences.

To determine the percent identity of two amino acids sequences or two nucleic acid sequences, the sequences are aligned for optimal comparison. For example, gaps can be introduced in the sequence of a first amino acid sequence or a first nucleic acid sequence for optimal alignment with the second amino acid sequence or second nucleic acid sequence. The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, the molecules are identical at that position.

The percent identity between the two sequences is a function of the number of identical positions shared by the sequences. Hence % identity = number of identical positions / total number of overlapping positions X 100.

In this comparison the sequences can be the same length or may be different in length. Optimal alignment of sequences for determining a comparison window may be conducted by the local homology algorithm of Smith and Waterman (*J. Theor. Biol.*, 91 (2) pgs. 370-380 (1981), by the homology alignment algorithm of Needleman and Wunsch, *J. Miol. Biol.*, 48(3) pgs. 443-453 (1972), by the search for similarity via the method of Pearson and Lipman, *PNAS, USA*, 85(5) pgs. 2444-2448 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA and TFASTA in the Wisconsin Genetics Software Package Release 7.0, Genetic Computer Group, 575, Science Drive, Madison, Wisconsin) or by inspection.

The best alignment (i.e., resulting in the highest percentage of identity over the comparison window) generated by the various methods is selected.

The term "sequence identity" means that two polynucleotide sequences are identical (i.e., on a nucleotide by nucleotide basis) over the window of comparison. The term "percentage of sequence identity" is calculated by comparing two optimally aligned sequences over the window of comparison, determining the number of positions at which the identical nucleic acid base (e.g., A, T, C, G, U, or I) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison (i.e., the window size) and multiplying the result by 100 to yield the percentage of sequence identity. The same process can be applied to polypeptide sequences.

The percentage of sequence identity of a nucleic acid sequence or an amino acid sequence can also be calculated using BLAST software (Version 2.06 of September 1998) with the default or user defined parameter.

The term "sequence similarity" means that amino acids can be modified while retaining the same function. It is known that amino acids are classified according to the nature of their side groups and some amino acids such as the basic amino acids can be interchanged for one another while their basic function is maintained.

The term "isolated" as used herein means that a biological material such as a nucleic acid or protein has been removed from its original environment in which it is naturally present. For example, a polynucleotide present in a plant, mammal or animal is present in its natural state and is not considered to be isolated. The same polynucleotide separated from the adjacent nucleic acid sequences in which it is naturally inserted in the genome of the plant or animal is considered as being "isolated."

The term "isolated" is not meant to exclude artificial or synthetic mixtures with other compounds, or the presence of impurities which do not interfere with the biological activity and which may be present, for example, due to incomplete purification, addition of stabilizers or mixtures with pharmaceutically acceptable excipients and the like.

"Isolated polypeptide" or "isolated protein" as used herein means a polypeptide or protein which is substantially free of those compounds that are normally associated with the polypeptide or protein in a naturally state such as other proteins or polypeptides, nucleic acids, carbohydrates, lipids and the like.

The term "purified" as used herein means at least one order of magnitude of purification is achieved, preferably two or three orders of magnitude, most preferably four or five orders of magnitude of purification of the starting material or of the natural material. Thus, the term "purified" as utilized herein does not mean that the material is 100% purified and thus excludes any other material.

The term "variants" when referring to, for example, polynucleotides encoding a polypeptide variant of a given reference polypeptide are polynucleotides that differ from the reference polypeptide but generally maintain their functional characteristics of the reference polypeptide. A variant of a polynucleotide may be a naturally occurring allelic variant or it may be a variant that is known naturally not to occur. Such non-naturally occurring variants of the reference polynucleotide can be made by, for example, mutagenesis techniques, including those mutagenesis techniques that are applied to polynucleotides, cells or organisms.

Generally, differences are limited so that the nucleotide sequences of the reference and variant are closely similar overall and, in many regions identical.

Variants of polynucleotides according to the present invention include, but are not limited to, nucleotide sequences which are at least 95% identical after alignment to the reference polynucleotide encoding the reference polypeptide. These variants can also have 96%, 97%, 98% and 99.999% sequence identity to the reference polynucleotide.

Nucleotide changes present in a variant polynucleotide may be silent, which means that these changes do not alter the amino acid sequences encoded by the reference polynucleotide.

Substitutions, additions and/or deletions can involve one or more nucleic acids. Alterations can produce conservative or non-conservative amino acid substitutions, deletions and/or additions.

Variants of a prey or a SID® polypeptide encoded by a variant polynucleotide can possess a higher affinity of binding and/or a higher specificity of binding to its protein or polypeptide counterpart, against which it has been initially selected. In another context, variants can also loose their ability to bind to their protein or polypeptide counterpart.

By "fragment of a polynucleotide" or "fragment of a SID® polynucleotide" is meant that fragments of these sequences have at least 12 consecutive nucleotides, or between 12 and 5,000 consecutive nucleotides, or between 12 and 10,000 consecutive nucleotides, or between 12 and 20,000 consecutive nucleotides.

By "fragment of a polypeptide" or "fragment of a SID® polypeptide" is meant that fragments of these sequences have at least 4 consecutive amino acids, or between 4 and 1,700 consecutive amino acids, or between 4 and 3,300 consecutive amino acids, or between 4 and 6,600 consecutive amino acids.

By "anabolic pathway" is meant a reaction or series of reactions in a metabolic pathway that synthesize complex molecules from simpler ones, usually requiring the input of energy. An anabolic pathway is the opposite of a catabolic pathway.

As used herein, a "catabolic pathway" is a series of reactions in a metabolic pathway that break down complex compounds into simpler ones, usually releasing energy in the process. A catabolic pathway is the opposite of an anabolic pathway.

As used herein, "drug metabolism" is meant the study of how drugs are processed and broken down by the body. Drug metabolism can involve the study of enzymes that break down drugs, the study of how different drugs interact within the body and how diet and other ingested compounds affect the way the body processes drugs.

As used herein, "metabolism" means the sum of all of the enzyme-catalyzed reactions in living cells that transform organic molecules.

By "secondary metabolism" is meant pathways producing specialized metabolic products that are not found in every cell.

As used herein, "SID®" means a Selected Interacting Domain and is identified as follows: for each bait polypeptide screened, selected prey polypeptides are compared. Overlapping fragments in the same ORF or CDS define the selected interacting domain.

As used herein the term "PIM®" means a protein-protein interaction map. This map is obtained from data acquired from a number of separate screens using different bait polypeptides and is designed to map out all of the interactions between the polypeptides.

The term "affinity of binding", as used herein, can be defined as the affinity constant Ka when a given SID® polypeptide of the present invention which binds to a polypeptide and is the following mathematical relationship:

$$Ka = \frac{[\text{SID®/polypeptide complex}]}{[\text{free SID®}]\,[\text{free polypeptide}]}$$

wherein [free SID®], [free polypeptide] and [SID®/polypeptide complex] consist of the concentrations at equilibrium respectively of the free SID® polypeptide, of the free polypeptide onto which the SID® polypeptide binds and of the complex formed between SID® polypeptide and the polypeptide onto which said SID® polypeptide specifically binds.

The affinity of a SID® polypeptide of the present invention or a variant thereof for its polypeptide counterpart can be assessed, for example, on a Biacore™ apparatus marketed by Amersham Pharmacia Biotech Company such as described by Szabo *et al.* (*Curr Opin Struct Biol* **5** pgs. 699-705 (1995)) and by Edwards and Leartherbarrow (*Anal. Biochem* 246 pgs. 1-6 (1997)).

As used herein the phrase "at least the same affinity" with respect to the binding affinity between a SID® polypeptide of the present invention to another polypeptide means that the Ka is identical or can be at least two-fold, at least three-fold or at least five fold greater than the Ka value of reference.

As used herein, the term "modulating compound" means a compound that inhibits or stimulates or can act on another protein which can inhibit or stimulate the protein-protein interaction of a complex of two polypeptides or the protein-protein interaction of two polypeptides.

More specifically, the present invention comprises complexes of polypeptides or polynucleotides encoding the polypeptides composed of a bait polypeptide, or a bait polynucleotide encoding a bait polypeptide and a prey polypeptide or a prey polynucleotide encoding a prey polypeptide. The prey polypeptide or prey polynucleotide encoding the prey polypeptide is capable of interacting with a bait polypeptide of interest in various hybrid systems.

As described in the background of the present invention, there are various methods known in the art to identify prey polypeptides that interact with bait polypeptides of interest. These methods include, but are not limited to, generic two-hybrid systems as described by Fields *et al.* (*Nature*, 340:245-246 (1989)) and more specifically in U.S. Patent Nos. 5,283,173, 5,468,614 and 5,667,973, which are hereby incorporated by reference; the

reverse two-hybrid system described by Vidal *et al.* (*supra*); the two plus one hybrid method described, for example, in Tirode *et al.* (*supra*); the yeast forward and reverse 'n'-hybrid systems as described in Vidal and Legrain (*supra*); the method described in WO 99/42612; those methods described in Legrain *et al.* (*FEBS Letters* 480 pgs. 32-36 (2000)) and the like.

The present invention is not limited to the type of method utilized to detect protein-protein interactions and therefore any method known in the art and variants thereof can be used. It is however better to use the method described in WO99/42612 or WO00/66722, both references incorporated herein by reference due to the methods' sensitivity, reproducibility and reliability.

Protein-protein interactions can also be detected using complementation assays such as those described by Pelletier *et al.* at http://www.abrf.org/JBT/Articles/JBT0012/jbt0012.html, WO 00/07038 and WO98/34120.

Although the above methods are described for applications in the yeast system, the present invention is not limited to detecting protein-protein interactions using yeast, but also includes similar methods that can be used in detecting protein-protein interactions in, for example, mammalian systems as described, for example in Takacs *et al.* (*Proc. Natl. Acad. Sci., USA*, **90** (21):10375-79 (1993)) and Vasavada *et al.* (*Proc. Natl. Acad. Sci., USA*, 88 (23):10686-90 (1991)), as well as a bacterial two-hybrid system as described in Karimova *et al.* (1998), WO99/28746, WO00/66722 and Legrain *et al.* (*FEBS Letters,* **480** pgs. 32-36 (2000)).

The above-described methods are limited to the use of yeast, mammalian cells and *Escherichia coli* cells, the present invention is not limited in this manner. Consequently, mammalian and typically human cells, as well as bacterial, yeast, fungus, insect, nematode and plant cells are encompassed by the present invention and may be transfected by the nucleic acid or recombinant vector as defined herein.

Examples of suitable cells include, but are not limited to, VERO cells, HELA cells such as ATCC No. CCL2, CHO cell lines such as ATCC No. CCL61, COS cells such as COS-7 cells and ATCC No. CRL 1650 cells, W138, BHK, HepG2, 3T3 such as ATCC No. CRL6361, A549, PC12, K562 cells, 293 cells, Sf9 cells such as ATCC No. CRL1711 and Cv1 cells such as ATCC No. CCL70.

Other suitable cells that can be used in the present invention include, but are not limited to, prokaryotic host cells strains such as *Escherichia coli,* (e.g., strain DH5-α), *Bacillus subtilis, Salmonella typhimurium*, or strains of the genera of *Pseudomonas, Streptomyces* and *Staphylococcus.*

Further suitable cells that can be used in the present invention include yeast cells such as those of *Saccharomyces* such as *Saccharomyces cerevisiae.*

The bait polynucleotide, as well as the prey polynucleotide can be prepared according to the methods known in the art such as those described above in the publications and patents reciting the known method *per se*.

The bait and the prey polynucleotide of the present invention is obtained from transforming growth factor β cDNA, or variants of cDNA fragment from a library of transforming growth factor β, and fragments from the genome or transcriptome of transforming growth factor β cDNA ranging from about 12 to about 5,000, or about 12 to about 10,000 or from about 12 to about 20,000. The prey polynucleotide is then selected, sequenced and identified.

A transforming growth factor β super-family of cytokines prey library is prepared from the transforming growth factor β cDNA and constructed in the specially designed prey vector pP6 as shown in Figure 3 after ligation of suitable linkers such that every cDNA insert is fused to a nucleotide sequence in the vector that encodes the transcription activation domain of a reporter gene. Any transcription activation domain can be used in the present invention. Examples include, but are not limited to, Gal4,YP16, B42, His and the like. Toxic reporter genes, such as CAT^R, CYH2, CYH1, URA3, bacterial and fungi toxins and the like can be used in reverse two-hybrid systems.

The polypeptides encoded by the nucleotide inserts of the transforming growth factor β prey library thus prepared are termed "prey polypeptides" in the context of the presently described selection method of the prey polynucleotides.

The bait polynucleotides can be inserted in bait plasmid pB27 or pB28 as illustrated in Figure 8 and Figure 9. The bait polynucleotide insert is fused to a polynucleotide encoding the binding domain of, for example, the Gal4 DNA binding domain and the shuttle expression vector is used to transform cells.

The bait polynucleotides used in the present invention are described in Table 1.

As stated above, any cells can be utilized in transforming the bait and prey polynucleotides of the present invention including mammalian cells, bacterial cells, yeast cells, insect cells and the like.

In an embodiment, the present invention identifies protein-protein interactions in yeast. In using known methods a prey positive clone is identified containing a vector which comprises a nucleic acid insert encoding a prey polypeptide which binds to a bait polypeptide of interest. The method in which protein-protein interactions are identified comprises the following steps:

i) mating at least one first haploid recombinant yeast cell clone from a recombinant yeast cell clone library that has been transformed with a plasmid containing the prey polynucleotide to be assayed with a second haploid recombinant yeast cell clone transformed with a plasmid containing a bait polynucleotide encoding for the bait polypeptide;

ii) cultivating diploid cell clones obtained in step i) on a selective medium; and

iii) selecting recombinant cell clones which grow on the selective medium.

This method may further comprise the step of:

iv) characterizing the prey polynucleotide contained in each recombinant cell clone which is selected in step iii).

In yet another embodiment of the present invention, *in lieu* of yeast, *Escherichia coli* is used in a bacterial two-hybrid system, which encompasses a similar principle to that described above for yeast, but does not involve mating for characterizing the prey polynucleotide.

In yet another embodiment of the present invention, mammalian cells and a method similar to that described above for yeast for characterizing the prey polynucleotide are used.

By performing the yeast, bacterial or mammalian two-hybrid system, it is possible to identify for one particular bait an interacting prey polypeptide. The prey polynucleotide that has been selected by testing the library of preys in a screen using the two-hybrid, two plus one hybrid methods and the like, encodes the polypeptide interacting with the protein of interest.

The present invention is also directed, in a general aspect, to a complex of polypeptides, polynucleotides encoding the polypeptides composed of a bait polypeptide or bait polynucleotide encoding the bait polypeptide and a prey polypeptide or prey polynucleotide encoding the prey polypeptide capable of interacting with the bait polypeptide of interest. These complexes are identified in Table 2.

In another aspect, the present invention relates to a complex of polynucleotides consisting of a first polynucleotide, or a fragment thereof, encoding a prey polypeptide that interacts with a bait polypeptide and a second polynucleotide or a fragment thereof. This fragment has at least 12 consecutive nucleotides, but can have between 12 and 5,000 consecutive nucleotides, or between 12 and 10,000 consecutive nucleotides or between 12 and 20,000 consecutive nucleotides.

The complexes of the two interacting polypeptides listed in Table 2 and the sets of two polynucleotides encoding these polypeptides also form part of the present invention.

In yet another embodiment, the present invention relates to an isolated complex of at least two polypeptides encoded by two polynucleotides wherein said two polypeptides are associated in the complex by affinity binding and are depicted in columns 1 and 4 of Table 2.

In yet another embodiment, the present invention relates to an isolated complex comprising at least a polypeptide as described in column 1 of Table 2 and a polypeptide as described in column 4 of Table 2. The present invention is not limited to these polypeptide complexes alone but also includes the isolated complex of the two polypeptides in which

fragments and/or homologous polypeptides exhibit at least 95% sequence identity, as well as from 96% sequence identity to 99.999% sequence identity.

Also encompassed in another embodiment of the present invention is an isolated complex in which the SID® of the prey polypeptides encoded by SEQ ID N°27 to 64 in Table 3 form the isolated complex.

Besides the isolated complexes described above, nucleic acids coding for a Selected Interacting Domain (SID®) polypeptide or a variant thereof or any of the nucleic acids set forth in Table 3 can be inserted into an expression vector which contains the necessary elements for the transcription and translation of the inserted protein-coding sequence. Such transcription elements include a regulatory region and a promoter. Thus, the nucleic acid which may encode a marker compound of the present invention is operably linked to a promoter in the expression vector. The expression vector may also include a replication origin.

A wide variety of host/expression vector combinations are employed in expressing the nucleic acids of the present invention. Useful expression vectors that can be used include, for example, segments of chromosomal, non-chromosomal and synthetic DNA sequences. Suitable vectors include, but are not limited to, derivatives of SV40 and pcDNA and known bacterial plasmids such as col EI, pCR1, pBR322, pMal-C2, pET, pGEX as described by Smith et al (1988), pMB9 and derivatives thereof, plasmids such as RP4, phage DNAs such as the numerous derivatives of phage I such as NM989, as well as other phage DNA such as M13 and filamentous single stranded phage DNA; yeast plasmids such as the 2 micron plasmid or derivatives of the 2m plasmid, as well as centomeric and integrative yeast shuttle vectors; vectors useful in eukaryotic cells such as vectors useful in insect or mammalian cells; vectors derived from combinations of plasmids and phage DNAs, such as plasmids that have been modified to employ phage DNA or the expression control sequences; and the like.

For example in a baculovirus expression system, both non-fusion transfer vectors, such as, but not limited to pVL941 (*Bam*HI cloning site Summers), pVL1393 (*Bam*HI, *Sma*I, *Xba*I, *Eco*RI, *Not*I, *Xma*III, *Bg*lII and *Pst*I cloning sites; Invitrogen), pVL1392 (*Bg*lII, *Pst*I, *Not*I, *Xma*III, *Eco*RI, *Xba*I, *Sma*I and *Bam*HI cloning site; Summers and Invitrogen) and pBlue*Bac*III (*Bam*HI, *Bg*lII, *Pst*I, *Nco*I and *Hind*III cloning site, with blue/white recombinant screening, Invitrogen), and fusion transfer vectors such as, but not limited to, pAc700 (*Bam*HI and *Kpn*I cloning sites, in which the *Bam*HI recognition site begins with the initiation codon; Summers), pAc701 and pAc70-2 (same as pAc700, with different reading frames), pAc360 (*Bam*HI cloning site 36 base pairs downstream of a polyhedrin initiation codon; Invitrogen (1995)) and pBlueBacHisA, B, C (three different reading frames with *Bam*HI, *Bg*lII, *Pst*I, *Nco*I and *Hind*III cloning site, an N-terminal peptide for ProBond purification and blue/white recombinant screening of plaques; Invitrogen (220) can be used.

Mammalian expression vectors contemplated for use in the invention include vectors with inducible promoters, such as the dihydrofolate reductase promoters, any expression vector with a DHFR expression cassette or a DHFR/methotrexate co-amplification vector such as pED (*Pst*I, *Sal*I, Sbal, Smal and *Eco*RI cloning sites, with the vector expressing both the cloned gene and DHFR; Kaufman, 1991). Alternatively a glutamine synthetase/methionine sulfoximine co-amplification vector, such as pEE14 (*Hind*III, *Xba*I, *Sma*I, *Sba*I, *Eco*RI and *Bcl*I cloning sites in which the vector expresses glutamine synthetase and the cloned gene; Celltech). A vector that directs episomal expression under the control of the Epstein Barr Virus (EBV) or nuclear antigen (EBNA) can be used such as pREP4 (*Bam*HI, *Sfi*I, *Xho*I, *Not*I, *Nhe*I, *Hind*III, *Nhe*I, *Pvu*II and *Kpn*I cloning sites, constitutive RSV-LTR promoter, hygromycin selectable marker; Invitrogen), pCEP4 (*Bam*HI, *Sfi*I, *Xho*I, *Not*I, *Nhe*I, *Hind*III, *Nhe*I, *Pvu*II and *Kpn*I cloning sites, constitutive hCMV immediate early gene promoter, hygromycin selectable marker; Invitrogen), pMEP4 (*Kpn*I, *Pvu*I, *Nhe*I, *Hind*III, *Not*I, *Xho*I, *Sfi*I, *Bam*HI cloning sites, inducible methallothionein IIa gene promoter, hygromycin selectable marker, Invitrogen), pREP8 (*Bam*HI, *Xho*I, *Not*I, *Hind*III, *Nhe*I and *Kpn*I cloning sites, RSV-LTR promoter, histidinol selectable marker; Invitrogen), pREP9 (*Kpn*I, *Nhe*I, *Hind*III, *Not*I, *Xho*I, *Sfi*I, *Bam*HI cloning sites, RSV-LTR promoter, G418 selectable marker; Invitrogen), and pEBVHis (RSV-LTR promoter, hygromycin selectable marker, N-terminal peptide purifiable via ProBond resin and cleaved by enterokinase; Invitrogen).

Selectable mammalian expression vectors for use in the invention include, but are not limited to, pRc/CMV (*Hind*III, *Bst*XI, *Not*I, *Sba*I and *Apa*I cloning sites, G418 selection, Invitrogen), pRc/RSV (*Hind*II, *Spe*I, *Bst*XI, *Not*I, *Xba*I cloning sites, G418 selection, Invitrogen) and the like. Vaccinia virus mammalian expression vectors (see, for example Kaufman 1991 that can be used in the present invention include, but are not limited to, pSC11 (*Sma*I cloning site, TK- and β-gal selection), pMJ601 (*Sal*I, *Sma*I, *Afl*II, *Nar*I, *Bsp*MII, *Bam*HI, *Apa*I, *Nhe*I, *Sac*II, *Kpn*I and *Hind*III cloning sites; TK- and β-gal selection), pTKgptF1S (*Eco*RI, *Pst*I, *Sal*II, *Acc*I, *Hind*II, *Sba*I, *Bam*HI and *Hpa* cloning sites, TK or XPRT selection) and the like.

Yeast expression systems that can also be used in the present include, but are not limited to, the non-fusion pYES2 vector (*Xba*I, *Sph*I, *Sho*I, *Not*I, *GstX*I, *Eco*RI, *Bst*XI, *Bam*HI, *Sac*I, *Kpn*I and *Hind*III cloning sites, Invitrogen), the fusion pYESHisA, B, C (*Xba*I, *Sph*I, *Sho*I, *Not*I, *Bst*XI, *Eco*RI, *Bam*HI, *Sac*I, *Kpn*I and *Hind*III cloning sites, N-terminal peptide purified with ProBond resin and cleaved with enterokinase; Invitrogen), pRS vectors and the like.

Consequently, mammalian and typically human cells, as well as bacterial, yeast, fungi, insect, nematode and plant cells an used in the present invention and may be transfected by the nucleic acid or recombinant vector as defined herein.

Examples of suitable cells include, but are not limited to, VERO cells, HELA cells such as ATCC No. CCL2, CHO cell lines such as ATCC No. CCL61, COS cells such as COS-7 cells and ATCC No. CRL 1650 cells, W138, BHK, HepG2, 3T3 such as ATCC No. CRL6361, A549, PC12, K562 cells, 293 cells, Sf9 cells such as ATCC No. CRL1711 and Cv1 cells such as ATCC No. CCL70.

Other suitable cells that can be used in the present invention include, but are not limited to, prokaryotic host cells strains such as *Escherichia coli*, (e.g., strain DH5-α), *Bacillus subtilis*, *Salmonella typhimurium*, or strains of the genera of *Pseudomonas*, *Streptomyces* and *Staphylococcus*.

Further suitable cells that can be used in the present invention include yeast cells such as those of *Saccharomyces* such as *Saccharomyces cerevisiae*.

Besides the specific isolated complexes, as described above, the present invention relates to and also encompasses SID® polynucleotides. As explained above, for each bait polypeptide, several prey polypeptides may be identified by comparing and selecting the intersection of every isolated fragment that are included in the same polypeptide. Thus the SID® polynucleotides of the present invention are represented by the shared nucleic acid sequences of SEQ ID N° 27 to 64 encoding the SID® polypeptides of SEQ ID N° 65 to 102 in columns 5 and 7 of Table 3, respectively.

The present invention is not limited to the SID® sequences as described in the above paragraph, but also includes fragments of these sequences having at least 12 consecutive nucleic acids, between 12 and 5,000 consecutive nucleic acids and between 12 and 10,000 consecutive nucleic acids and between 12 and 20,000 consecutive nucleic acids, as well as variants thereof. The fragments or variants of the SID® sequences possess at least the same affinity of binding to its protein or polypeptide counterpart, against which it has been initially selected. Moreover this variant and/or fragments of the SID® sequences alternatively can have between 95% and 99.999% sequence identity to its protein or polypeptide counterpart.

According to the present invention variants of polynucleotide or polypeptides can be created by known mutagenesis techniques either *in vitro* or *in vivo*. Such a variant can be created such that it has altered binding characteristics with respect to the target protein and more specifically that the variant binds the target sequence with either higher or lower affinity.

Polynucleotides that are complementary to the above sequences which include the polynucleotides of the SID®'s, their fragments, variants and those that have specific sequence identity are also included in the present invention.

The polynucleotide encoding the SID® polypeptide, fragment or variant thereof can also be inserted into recombinant vectors which are described in detail above.

The present invention also relates to a composition comprising the above-mentioned recombinant vectors containing the SID® polynucleotides in Table 3, fragments or variants thereof, as well as recombinant host cells transformed by the vectors. The recombinant host cells that can be used in the present invention were discussed in greater detail above.

The compositions comprising the recombinant vectors can contain physiological acceptable carriers such as diluents, adjuvants, excipients and any vehicle in which this composition can be delivered therapeutically and can include, but is are not limited to sterile liquids such as water and oils.

In yet another embodiment, the present invention relates to a method of selecting modulating compounds, as well as the modulating molecules or compounds themselves which may be used in a pharmaceutical composition. These modulating compounds may act as a cofactor, as an inhibitor, as antibodies, as tags, as a competitive inhibitor, as an activator or alternatively have agonistic or antagonistic activity on the protein-protein interactions.

The activity of the modulating compound does not necessarily, for example, have to be 100% activation or inhibition. Indeed, even partial activation or inhibition can be achieved that is of pharmaceutical interest.

The modulating compound can be selected according to a method which comprises:

(a) cultivating a recombinant host cell with a modulating compound on a selective medium and a reporter gene the expression of which is toxic for said recombinant host cell wherein said recombinant host cell is transformed with two vectors:

(i) wherein said first vector comprises a polynucleotide encoding a first hybrid polypeptide having a DNA binding domain;

(ii)wherein said second vector comprises a polynucleotide encoding a second hybrid polypeptide having a transcriptional activating domain that activates said toxic reporter gene when the first and second hybrid polypeptides interact;

(b) selecting said modulating compound which inhibits or permits the growth of said recombinant host cell.

Thus, the present invention relates to a modulating compound that inhibits the protein-protein interactions of a complex of two polypeptides of columns 1 and 4 of Table 2. The

present invention also relates to a modulating compound that activates the protein-protein interactions of a complex of two polypeptides of columns 1 and 4 of Table 2.

In yet another embodiment, the present invention relates to a method of selecting a modulating compound, which modulating compound inhibits the interactions of two polypeptides of columns 1 and 4 of Table 2. This method comprises:

(a) cultivating a recombinant host cell with a modulating compound on a selective medium and a reporter gene the expression of which is toxic for said recombinant host cell wherein said recombinant host cell is transformed with two vectors:

(i) wherein said first vector comprises a polynucleotide encoding a first hybrid polypeptide having a first domain of an enzyme;

(ii) wherein said second vector comprises a polynucleotide encoding a second hybrid polypeptide having an enzymatic transcriptional activating domain that activates said toxic reporter gene when the first and second hybrid polypeptides interact;

(b) selecting said modulating compound which inhibits or permits the growth of said recombinant host cell.

In the two methods described above any toxic reporter gene can be utilized including those reporter genes that can be used for negative selection including the URA3 gene, the CYH1 gene, the CYH2 gene and the like.

In yet another embodiment, the present invention provides a kit for screening a modulating compound. This kit comprises a recombinant host cell which comprises a reporter gene the expression of which is toxic for the recombinant host cell. The host cell is transformed with two vectors. The first vector comprises a polynucleotide encoding a first hybrid polypeptide having a DNA binding domain; and the second vector comprises a polynucleotide encoding a second hybrid polypeptide having a transcriptional activating domain that activates said toxic reporter gene when the first and second hybrid polypeptides interact.

In yet another embodiment, a kit is provided for screening a modulating compound by providing a recombinant host cell, as described in the paragraph above, but instead of a DNA binding domain, the first vector encodes a first hybrid polypeptide containing a first domain of a protein. The second vector encodes a second polypeptide containing a second part of a complementary domain of a protein that activates the toxic reporter gene when the first and second hybrid polypeptides interact.

In the selection methods described above, the activating domain can be p42 Gal 4, YP16 (HSV) and the DNA-binding domain can be derived from Gal4 or Lex A. The protein or enzyme can be adenylate cyclase, guanylate cyclase, DHFR and the like.

Examples of modulating compounds are set forth in Table 3.

In yet another embodiment, the present invention relates to a pharmaceutical composition comprising the modulating compounds for preventing or treating disorders and/or diseases involving members of the TGFβ family of cytokines in a human or animal, most preferably in a mammal.

This pharmaceutical composition comprises a pharmaceutically acceptable amount of the modulating compound. The pharmaceutically acceptable amount can be estimated from cell culture assays. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes or encompasses a concentration point or range having the desired effect in an *in vitro* system. This information can thus be used to accurately determine the doses in other mammals, including humans and animals.

The therapeutically effective dose refers to that amount of the compound that results in amelioration of symptoms in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or in experimental animals. For example, the LD50 (the dose lethal to 50% of the population) as well as the ED50 (the dose therapeutically effective in 50% of the population) can be determined using methods known in the art. The dose ratio between toxic and therapeutic effects is the therapeutic index which can be expressed as the ratio between LD 50 and ED50 compounds that exhibit high therapeutic indexes.

The data obtained from the cell culture and animal studies can be used in formulating a range of dosage of such compounds which lies preferably within a range of circulating concentrations that include the ED50 with little or no toxicity.

The pharmaceutical composition can be administered via any route such as locally, orally, systemically, intravenously, intramuscularly, mucosally, using a patch and can be encapsulated in liposomes, microparticles, microcapsules, and the like. The pharmaceutical composition can be embedded in liposomes or even encapsulated.

Any pharmaceutically acceptable carrier or adjuvant can be used in the pharmaceutical composition. The modulating compound will be preferably in a soluble form combined with a pharmaceutically acceptable carrier. The techniques for formulating and administering these compounds can be found in "*Remington's Pharmaceutical Sciences*" Mack Publication Co., Easton, PA, latest edition.

The mode of administration optimum dosages and galenic forms can be determined by the criteria known in the art taken into account the seriousness of the general condition of the mammal, the tolerance of the treatment and the side effects.

The present invention also relates to a method of treating or preventing diseases involving the trasduction pathways of members of the transforming growth factor β super-family of cytokines in a human or mammal in need of such treatment. This method comprises administering to a mammal in need of such treatment a pharmaceutically effective

amount of a modulating compound which binds to a targeted mammalian or human or inner ear cell protein. In a preferred embodiment, the modulating compound is a polynucleotide which may be placed under the control of a regulatory sequence which is functional in the mammal or human.

In yet another embodiment, the present invention relates to a pharmaceutical composition comprising a SID® polypeptide, a fragment or variant thereof. The SID® polypeptide, fragment or variant thereof can be used in a pharmaceutical composition provided that it is endowed with highly specific binding properties to a bait polypeptide of interest.

The original properties of the SID® polypeptide or variants thereof interfere with the naturally occurring interaction between a first protein and a second protein within the cells of the organism. Thus, the SID® polypeptide binds specifically to either the first polypeptide or the second polypeptide.

Therefore, the SID® polypeptides of the present invention or variants thereof interfere with protein-protein interactions between mammalian and especially human protein.

Thus, the present invention relates to a pharmaceutical composition comprising a pharmaceutically acceptable amount of a SID® polypeptide or variant thereof, provided that the variant has the above-mentioned two characteristics; i.e., that it is endowed with highly specific binding properties to a bait polypeptide of interest and is devoid of biological activity of the naturally occurring protein.

In yet another embodiment, the present invention relates to a pharmaceutical composition comprising a pharmaceutically effective amount of a polynucleotide encoding a SID® polypeptide or a variant thereof wherein the polynucleotide is placed under the control of an appropriate regulatory sequence. Appropriate regulatory sequences that are used are polynucleotide sequences derived from promoter elements and the like.

Polynucleotides that can be used in the pharmaceutical composition of the present invention include the nucleotide sequences of SEQ ID N° 27 to 64.

Besides the SID® polypeptides and polynucleotides, the pharmaceutical composition of the present invention can also include a recombinant expression vector comprising the polynucleotide encoding the SID® polypeptide, fragment or variant thereof.

The above described pharmaceutical compositions can be administered by any route such as orally, systemically, intravenously, intramuscularly, intradermally, mucosally, encapsulated, using a patch and the like. Any pharmaceutically acceptable carrier or adjuvant can be used in this pharmaceutical composition.

The SID® polypeptides as active ingredients will be preferably in a soluble form combined with a pharmaceutically acceptable carrier. The techniques for formulating and

administering these compounds can be found in "*Remington's Pharmaceutical Sciences*" *supra*.

The amount of pharmaceutically acceptable SID® polypeptides can be determined as described above for the modulating compounds using cell culture and animal models.

Such compounds can be used in a pharmaceutical composition to treat or prevent transforming growth factor β-mediated disorders and/or diseases.

Thus, the present invention also relates to a method of preventing or treating transforming growth factor β-mediated disorders and/or diseases in a mammal said method comprising the steps of administering to a mammal in need of such treatment a pharmaceutically effective amount of:

(1) a SID® polypeptide of SEQ ID N°65 to 105 or a variant thereof which binds to a targeted mammalian or typically human protein; or

(2) or SID® polynucleotide encoding a SID® polypeptide of SEQ ID N° 65 to 102 or a variant or a fragment thereof wherein said polynucleotide is placed under the control of a regulatory sequence which is functional in said mammal.

In another embodiment the present invention nucleic acids comprising a sequence of SEQ ID N° 27 to 64 which encodes the protein of sequence SEQ ID N° 65 to 102 and/or functional derivatives thereof are administered to modulate complex (from Table 2) function by way of gene therapy. Any of the methodologies relating to gene therapy available within the art may be used in the practice of the present invention such as those described by Goldspiel et al *Clin. Pharm.* **12** pgs. 488-505 (1993).

Delivery of the therapeutic nucleic acid into a patient may be direct *in vivo* gene therapy (i.e., the patient is directly exposed to the nucleic acid or nucleic acid-containing vector) or indirect *ex vivo* gene therapy (i.e., cells are first transformed with the nucleic acid in vitro and then transplanted into the patient).

For example for *in vivo* gene therapy, an expression vector containing the nucleic acid is administered in such a manner that it becomes intracellular; i.e., by infection using a defective or attenuated retroviral or other viral vectors as described, for example in U.S. Patent 4,980,286 or by Robbins et al, Pharmacol. *Ther.* , **80** No. 1 pgs. 35-47 (1998).

The various retroviral vectors that are known in the art are such as those described in Miller *et al.* (*Meth. Enzymol.* **217** pgs. 581-599 (1993)) which have been modified to delete those retroviral sequences which are not required for packaging of the viral genome and subsequent integration into host cell DNA. Also adenoviral vectors can be used which are advantageous due to their ability to infect non-dividing cells and such high-capacity adenoviral vectors are described in Kochanek (*Human Gene Therapy*, **10**, pgs. 2451-2459 (1999)). Chimeric viral vectors that can be used are those described by Reynolds *et al.*

(*Molecular Medecine Today*, pgs. 25 –31 (1999)). Hybrid vectors can also be used and are described by Jacoby *et al.* (*Gene Therapy*, **4**, pgs. 1282-1283 (1997)).

Direct injection of naked DNA or through the use of microparticle bombardment (e.g., Gene Gun®; Biolistic, Dupont) or by coating it with lipids can also be used in gene therapy. Cell-surface receptors/transfecting agents or through encapsulation in liposomes, microparticles or microcapsules or by administering the nucleic acid in linkage to a peptide which is known to enter the nucleus or by administering it in linkage to a ligand predisposed to receptor-mediated endocytosis (See Wu & Wu, J. Biol. Chem., 262 pgs. 4429-4432 (1987)) can be used to target cell types which specifically express the receptors of interest.

In another embodiment a nucleic acid ligand compound may be produced in which the ligand comprises a fusogenic viral peptide designed so as to disrupt endosomes, thus allowing the nucleic acid to avoid subsequent lysosomal degradation. The nucleic acid may be targeted *in vivo* for cell specific endocytosis and expression by targeting a specific receptor such as that described in WO92/06180, WO93/14188 and WO 93/20221. Alternatively the nucleic acid may be introduced intracellularly and incorporated within the host cell genome for expression by homologous recombination (See Zijlstra et al, *Nature*, **342**, pgs. 435-428 (1989)).

In *ex vivo* gene therapy, a gene is transferred into cells *in vitro* using tissue culture and the cells are delivered to the patient by various methods such as injecting subcutaneously, application of the cells into a skin graft and the intravenous injection of recombinant blood cells such as hematopoietic stem or progenitor cells.

Cells into which a nucleic acid can be introduced for the purposes of gene therapy include, for example, epithelial cells, endothelial cells, keratinocytes, fibroblasts, muscle cells, hepatocytes and blood cells. The blood cells that can be used include, for example, T-lymphocytes, B-lymphocytes, monocytes, macrophages, neutrophils, eosinophils, megakaryotcytes, granulocytes, hematopoietic cells or progenitor cells and the like.

In yet another embodiment the present invention relates to protein chips or protein microarrays. It is well known in the art that microarrays can contain more than 10,000 spots of a protein that can be robotically deposited on a surface of a glass slide or nylon filter. The proteins attach covalently to the slide surface, yet retain their ability to interact with other proteins or small molecules in solution. In some instances the protein samples can be made to adhere to glass slides by coating the slides with an aldehyde-containing reagent that attaches to primary amines. A process for creating microarrays is described, for example by MacBeath and Schreiber (*Science*, Volume 289, Number 5485, pgs, 1760-1763 (2000)) or (Service, *Science*, Vol, 289, Number 5485 pg. 1673 (2000)). An apparatus for controlling, dispensing and measuring small quantities of fluid is described, for example, in U.S. Patent No. 6,112,605.

The present invention also provides a record of protein-protein interactions, PIM®'s and any data encompassed in the following Tables. It will be appreciated that this record can be provided in paper or electronic or digital form.

The present invention also relates to the use of a SID® or an interaction or a prey to screen molecules that inhibit TGFβ or a TGFβ super-family of cytokines pathway, as well as molecules that inhibit TGFβ or a TGFβ super-family of cytokines pathway obtained by this screening method. The screening can occur in mammalian or yeast cells. Furthermore, the inhibition can be detected by fluorescence polarization, FRET, BRET, filter binding assays or radioactive techniques.

In order to fully illustrate the present invention and advantages thereof, the following specific examples are given, it being understood that the same are intended only as illustrative and in nowise limitative.

EXAMPLES

EXAMPLE 1: Preparation of a collection of random-primed cDNA fragments

1.A. Collection preparation and transformation in Escherichia coli

1.A.1. Random-primed cDNA fragment preparation

For mRNA sample from transforming growth factor β, random-primed cDNA was prepared from 5 µg of polyA+ mRNA using a TimeSaver cDNA Synthesis Kit (Amersham Pharmacia Biotech) and with 5 µg of random N9-mers according to the manufacturer's instructions. Following phenolic extraction, the cDNA was precipitated and resuspended in water. The resuspended cDNA was phosphorylated by incubating in the presence of T4 DNA Kinase (Biolabs) and ATP for 30 minutes at 37°C. The resulting phosphorylated cDNA was then purified over a separation column (Chromaspin TE 400, Clontech), according to the manufacturer's protocol.

1.A.2. Ligation of linkers to blunt-ended cDNA

Oligonucleotide HGX931 (5' end phosphorylated) 1 µg/µl and HGX932 1µg/µl were used.

Sequence of the oligo HGX931: 5'-GGGCCACGAA-3' (SEQ ID No.103)

Sequence of the oligo HGX932: 5'-TTCGTGGCCCCTG-3' (SEQ ID No.104)

Linkers were preincubated (5 minutes at 95°C, 10 minutes at 68°C, 15 minutes at 42°C) then cooled down at room temperature and ligated with cDNA fragments at 16°C overnight.

Linkers were removed on a separation column (Chromaspin TE 400, Clontech), according to the manufacturer's protocol.

1.A.3. Vector preparation

Plasmid pP6 (see Figure 3) was prepared by replacing the *SpelIXhoI* fragment of pGAD3S2X with the double-stranded oligonucleotide:

5'CTAGCCATGGCCGCAGGGGCCGCGGCCGCACTAGTGGGGATCCTTAATTAAGGGCC ACTGGGGCCCCC3' (SEQ ID No.105)

5'TCGAGGGGGCCCCAGTGGCCCTTAATTAAGGATCCCCACTAGTGCGGCCGCGGCCC CTGCGGCCATGG3' (SEQ ID No.106)

The pP6 vector was successively digested with *Sfi*1 and *Bam*HI restriction enzymes (Biolabs) for 1 hour at 37°C, extracted, precipitated and resuspended in water. Digested plasmid vector backbones were purified on a separation column (Chromaspin TE 400, Clontech), according to the manufacturer's protocol.

1.A.4. Ligation between vector and insert of cDNA

The prepared vector was ligated overnight at 15°C with the blunt-ended cDNA described in section 2 using T4 DNA ligase (Biolabs). The DNA was then precipitated and resuspended in water.

1.A.5. Library transformation in *Escherichia coli*

The DNA from section 1.A.4 was transformed into Electromax DH10B electrocompetent cells (Gibco BRL) with a Cell Porator apparatus (Gibco BRL). 1 ml SOC medium was added and the transformed cells were incubated at 37°C for 1 hour. 9 mls of SOC medium per tube was added and the cells were plated on LB+ampicillin medium. The colonies were scraped with liquid LB medium, aliquoted and frozen at -80°C.

1.B. Collection transformation in Saccharomyces cerevisiae

The *Saccharomyces cerevisiae* strain (YHGX13 (MATα Gal4Δ Gal80Δ ade2-101::KAN^R, his3, leu2-3, -112, trp1-901, ura3-52 URA3::UASGAL1-LacZ, Met)) was transformed with the cDNA library.

The plasmid DNA contained in *E. coli* were extracted (Qiagen) from aliquoted *E. coli* frozen cells (1.A.5.). *Saccharomyces cerevisiae* yeast YHGX13 in YPGlu were grown.

Yeast transformation was performed according to standard protocol (Giest *et al.* Yeast, 11, 355-360, 1995) using yeast carrier DNA (Clontech). This experiment leads to 10^4 to 5 x 10^4 cells/µg DNA. 2 x 10^4 cells were spread on DO-Leu medium per plate. The cells were aliquoted into vials containing 1 ml of cells and frozen at -80°C.

1.C. Construction of bait plasmids

For fusions of the bait protein to the DNA-binding domain of the GAL4 protein of *S. cerevisiae*, bait fragments were cloned into plasmid pB27and pB28.

Plasmid pB27 was prepared by replacing the ampicillin resistance of pB20 with the tetracyclin resistance.

MCS sequence EcoRI/PstI:

5'
AATTCGGGGCCGGACGGGCCGCGGCCGCACTAGTGGGGATCCTTAATTAAGGGCCAC
TGGGGCCCCTCGACCTGCA 3' (SEQ ID No 107)

5'
GGTCGAGGGGCCCCAGTGGCCCTTAATTAAGGATCCCCACTAGTGCGGCCGCGGCCC
GTCCGGCCCCG 3' (SEQ ID No 108)

Plasmid pB28 was prepared by replacing the EcoRI/PstI polylinker fragment of pB27 with the double stranded DNA fragment :

5'GAATTCGGGGCCGCAGGGGCCGCGGCCGCACTAGTGGGGATCCTTAATTAAGGGCC
ACTGGGGCCCCTCGACCTGCAG 3' (SEQ ID No 109)

5'CTGCAGGTCGAGGGGCCCCAGTGGCCCTTAATTAAGGATCCCCACTAGTGCGGCCG
CGGCCCCTGCGGCCCCGAATTC 3'(SEQ ID No 110)

The amplification of the bait ORF was obtained by PCR using the Pfu proof-reading *Taq* polymerase (Stratagene), 10 pmol of each specific amplification primer and 200 ng of plasmid DNA as template.

The PCR program was set up as follows :

94°	45"		
94°	45"	}	x 30 cycles
48°	45"		
72°	6'		
72°	10'		
15°	∞		

The amplification was checked by agarose gel electrophoresis.

The PCR fragments were purified with Qiaquick column (Qiagen) according to the manufacturer's protocol.

Purified PCR fragments were digested with adequate restriction enzymes.

The PCR fragments were purified with Qiaquick column (Qiagen) according to the manufacturer's protocol.

The digested PCR fragments were ligated into an adequately digested and dephosphorylated bait vector (pB27 or pB28) according to standard protocol (Sambrook *et al.*) and were transformed into competent bacterial cells. The cells were grown, the DNA extracted and the plasmid was sequenced.

Example 2 : Screening the collection with the two-hybrid in yeast system

2.A. The mating protocol

The mating two-hybrid in yeast system (as described by Legrain *et al.*, *Nature Genetics*, vol. 16, 277-282 (1997), *Toward a functional analysis of the yeast genome through*

exhaustive two-hybrid screens) was used for its advantages but one could also screen the cDNA collection in classical two-hybrid system as described in Fields *et al.* or in a yeast reverse two-hybrid system.

The mating procedure allows a direct selection on selective plates because the two fusion proteins are already produced in the parental cells. No replica plating is required.

This protocol was written for the use of the library transformed into the YHGX13 strain.

For bait proteins fused to the DNA-binding domain of GAL4, bait-encoding plasmids were first transformed into *S. cerevisiae* (CG1945 strain (MATa Gal4-542 Gal180-538 ade2-101 his3Δ200, leu2-3,112, trp1-901, ura3-52, lys2-801, URA3::GAL4 17mers (X3)-CyC1TATA-LacZ, LYS2::GAL1UAS-GAL1TATA-HIS3 CYH^R)) according to step 1.B. and spread on DO-Trp medium.

For bait proteins fused to the DNA-binding domain of LexA, bait-encoding plasmids were first transformed into *S. cerevisiae* (L40Δgal4 strain (MATa ade2, trp1-901, leu2 3,112, lys2-801, his3Δ200, LYS2::$(lexAop)_4$-HIS3, ura3-52::URA3 $(lexAop)_8$-LacZ, GAL4::Kan^R)) according to step 1.B. and spread on DO-Trp medium.

Day 1, morning : preculture

The cells carrying the bait plasmid obtained at step 1.C. were precultured in 20 ml DO-Trp medium and grown at 30°C with vigorous agitation.

Day 1, late afternoon : culture

The OD_{600nm} of the DO-Trp pre-culture of cells carrying the bait plasmid was measured. The OD_{600nm} must lie between 0.1 and 0.5 in order to correspond to a linear measurement.

50 ml DO-Trp at OD_{600nm} 0.006/ml was inoculated and grown overnight at 30°C with vigorous agitation.

Day 2 : mating

medium and plates

2 YPGlu 15cm plates

50 ml tube with 13 ml DO-Leu-Trp-His

100 ml flask with 5 ml of YPGlu

8 DO-Leu-Trp-His plates

2 DO-Leu-Trp plates

The OD_{600nm} of the DO-Trp culture was measured. It should be around 1.

For the mating, twice as many bait cells as library cells were used. To get a good mating efficiency, one must collect the cells at 10^8 cells per cm^2.

The amount of bait culture (in ml) that makes up 50 OD_{600nm} units for the mating with the prey library was estimated.

A vial containing the library of step 1B was thawed slowly on ice. 1.0ml of the vial was added to 20 ml YPGlu. Those cells were recovered at 30°C, under gentle agitation for 10 minutes.

Mating

The 50 OD_{600nm} units of bait culture was placed into a 50 ml falcon tube.

The library of step 1B culture was added to the bait culture, then centrifuged, the supernatant discarded and resuspended in 1.6ml YPGlu medium.

The cells were distributed onto two 15cm YPGlu plates with glass beads. The cells were spread by shaking the plates. The plate cells-up at 30°C for 4h30min were incubated.

Collection of mated cells

The plates were washed and rinsed with 6ml and 7ml respectively of DO-Leu-Trp-His. Two parallel serial ten-fold dilutions were performed in 500µl DO-Leu-Trp-His up to 1/10,000. 50µl of each 1/1,000 dilution was spread onto DO-Leu-Trp plates. 22.4ml of collected cells were spread in 400µl aliquots on DO-Leu-Trp-His+Tet plates.

Day 4

Clones that were able to grow on DO-Leu-Trp-His+Tetracyclin were then selected. This medium allows one to isolate diploid clones presenting an interaction.

The His+ colonies were counted on control plates.

The number of His+ cell clones will define which protocol is to be processed :

Upon 60.10^6 Trp+Leu+ colonies :

- if the number His+ cell clones <285: then use the process stamp overlay protocol on all colonies
- if the number of His+ cell clones >285 and <5000: then process via overlay and then stamp overlay protocols on blue colonies (2.B and 2.C).
- if number of His+ cell clones >5000: repeat screen using DO-Leu-Trp-His+Tetracyclin plates containing 3-aminotriazol.

2.B. The X-Gal overlay assay

The X-Gal overlay assay was performed directly on the selective medium plates after scoring the number of His^+ colonies.

Materials

A waterbath was set up. The water temperature should be 50°C.

- 0.5 M Na_2HPO_4 pH 7.5.
- 1.2% Bacto-agar.
- 2% X-Gal in DMF.
- Overlay mixture : 0.25 M Na_2HPO_4 pH7.5, 0.5% agar, 0.1% SDS, 7% DMF (LABOSI), 0.04% X-Gal (ICN). For each plate, 10 ml overlay mixture are needed.
- DO-Leu-Trp-His plates.

- Sterile toothpicks.

Experiment

The temperature of the overlay mix should be between 45°C and 50°C. The overlay-mix was poured over the plates in portions of 10 ml. When the top layer was settled, they were collected. The plates were incubated overlay-up at 30°C and the time was noted. Blue colonies were checked for regularly. If no blue colony appeared, overnight incubation was performed. Using a pen the number of positives was marked. The positives colonies were streaked on fresh DO-Leu-Trp-His plates with a sterile toothpick.

2.C. The stamp overlay assay

His+ colonies were grown overnight at 30°C in microtiter plates containing DO-Leu-Trp-His+Tetracyclin medium with shaking. The day after the overnight culture, the 96 colonies were stamped on a 15cm plate of DO-Leu-Trp-His. 4 control yeast colonies were spotted on the same plate. After 2 days of growing at 30°C, an overlay assay was performed on this plate with 80ml of overlay mixture (see step 2.B.). After 2 hours of incubation, the plate was photographed with a CCD camera. The blue intensity was quantified by Genetools® software (SYNGENE) and normalized to the control spots.

Example 3 : Identification of positive clones

3.A. PCR on yeast colonies

Introduction

PCR amplification of fragments of plasmid DNA directly on yeast colonies is a quick and efficient procedure to identify sequences cloned into this plasmid. It is directly derived from a published protocol (Wang H. et al., *Analytical Biochemistry,* **237**, 145-146, (1996)). However, it is not a standardized protocol and it varies from strain to strain and it is dependent of experimental conditions (number of cells, *Taq* polymerase source, etc). This protocol should be optimized to specific local conditions.

Materials

- For 1 well, PCR mix composition was:

  32.5 µl water,

  5 µl 10X PCR buffer (Pharmacia),

  1 µl dNTP 10 mM,

  0.5 µl *Taq* polymerase (5u/µl) (Pharmacia),

  0.5 µl oligonucleotide ABS1 10 pmole/µl: 5'-GCGTTTGGAATCACTACAGG-3' (SEQ ID No.111)

  0.5 µl oligonucleotide ABS2 10 pmole/µl: 5'-CACGATGCACGTTGAAGTG-3' (SEQ ID No.112)

- 1 N NaOH.

Experiment

The positive colonies were grown overnight at 30°C on a 96 well cell culture cluster (Costar), containing 150 µl DO-Leu-Trp-His+Tetracyclin with shaking. The culture was resuspended and 100 µl was transferred immediately on a Thermowell 96 (Costar) and centrifuged for 5 minutes at 4,000 rpm at room temperature. The supernatant was removed. 5 µl NaOH was added to each well and shaken for 1 minute.

The Thermowell was placed in the thermocycler (GeneAmp 9700, Perkin Elmer) for 5 minutes at 99.9°C and then 10 minutes at 4°C. In each well, the PCR mix was added and shaken well.

The PCR program was set up as followed:

94°C	3 minutes	
94°C	30 seconds	x 35 cycles
53°C	1 minute 30 seconds	
72°C	3 minutes	
72°C	5 minutes	
15°C	∞	

The quality, the quantity and the length of the PCR fragment was checked on an agarose gel. The length of the cloned fragment was the estimated length of the PCR fragment minus 300 base pairs that corresponded to the amplified flanking plasmid sequences.

3.B. Plasmids rescue from yeast by electroporation

Introduction

The previous protocol of PCR on yeast cell may not be successful, in such a case, plasmids from yeast by electroporation can be rescued. This experiment allows the recovery of prey plasmids from yeast cells by transformation of *E. coli* with a yeast cellular extract. The prey plasmid can then be amplified and the cloned fragment can be sequenced.

Materials

Plasmid rescue

Glass beads 425-600 µm (Sigma)

Phenol/chloroform (1/1) premixed with isoamyl alcohol (Amresco)

Extraction buffer : 2% Triton X100, 1% SDS, 100 mM NaCl, 10 mM TrisHCl pH 8.0, 1 mM EDTA pH 8.0.

Mix ethanol/NH_4Ac : 6 volumes ethanol with 7.5 M NH_4 Acetate, 70% Ethanol and yeast cells in patches on plates.

Electroporation

SOC medium

M9 medium

Selective plates : M9-Leu+Ampicillin

2 mm electroporation cuvettes (Eurogentech)

Experiment

Plasmid rescue

The cell patch on DO-Leu-Trp-His was prepared with the cell culture of section 2.C. The cell of each patch was scraped into an Eppendorf tube, 300 µl of glass beads was added in each tube, then, 200 µl extraction buffer and 200 µl phenol:chloroform:isoamyl alcohol (25:24:1) was added.

The tubes were centrifuged for 10 minutes at 15,000 rpm.

180 µl supernatant was transferred to a sterile Eppendorf tube and 500 µl each of ethanol/NH_4Ac was added and the tubes were vortexed. The tubes were centrifuged for 15 minutes at 15,000 rpm at 4°C. The pellet was washed with 200 µl 70% ethanol and the ethanol was removed and the pellet was dried. The pellet was resuspended in 10 µl water. Extracts were stored at -20°C.

Electroporation

Materials: Electrocompetent MC1066 cells prepared according to standard protocols (Sambrook et al. *supra*).

1 µl of yeast plasmid DNA-extract was added to a pre-chilled Eppendorf tube, and kept on ice.

1 µl plasmid yeast DNA-extract sample was mixed and 20 µl electrocompetent cells was added and transferred in a cold electroporation cuvette.

The Biorad electroporator was set on 200 ohms resistance, 25 µF capacity; 2.5 kV. The cuvette was placed in the cuvette holder and electroporation was performed.

1 ml of SOC was added into the cuvette and the cell-mix was transferred into a sterile Eppendorf tube. The cells were recovered for 30 minutes at 37°C, then spun down for 1 minute at 4,000 x g and the supernatant was poured off. About 100 µl medium was kept and used to resuspend the cells and spread them on selective plates (e.g., M9-Leu plates). The plates were then incubated for 36 hours at 37°C.

One colony was grown and the plasmids were extracted. The presence and the size of the insert were checked for through enzymatic digestion and agarose gel electrophoresis. The insert was then sequenced.

Example 4 : Protein-protein interaction

For each bait, the previous protocol leads to the identification of prey polynucleotide sequences. Using a suitable software program (e.g., Blastwun, available on the Internet site of the University of Washington: http://bioweb.pasteur.fr/seganal/interfaces/blastwu.html), the mRNA transcript that is encoded by the prey fragment may be identified and whether the

fusion protein encoded is in the same open reading frame of translation as the predicted protein or not can be determined.

Alternatively, prey nucleotide sequences can be compared with one another and those which share identity over a significant region (60nt) can be grouped together to form a contiguous sequence (Contig) whose identity can be ascertained in the same manner as for individual prey fragments described above.

Example 5 : Identification of SID®

By comparing and selecting the intersection of all isolated fragments that are included in the same polypeptide, one can define the Selected Interacting Domain (SID®) is determined as illustrated in Figure 6. The SID® is illustrated in Table 3.

Example 6: Making of polyclonal and monoclonal antibodies

The protein-protein complex of columns 1 and 4 of Table 2 is injected into mice and polyclonal and monoclonal antibodies are made following the procedure set forth in Sambrook et al *supra*.

More specifically, mice are immunized with an immunogen comprising the above mentionned complexes conjugated to keyhole limpet hemocyanin using glutaraldehyde or EDC as is well known in the art. The complexes can also be stabilized by crosslinking as described in WO 00/37483. The immunogen is then mixed with an adjuvant. Each mouse receives four injections of 10 µg to 100 µg of immunogen, and after the fourth injection, blood samples are taken from the mice to determine if the serum contains antibodies to the immunogen. Serum titer is determined by ELISA or RIA. Mice with sera indicating the presence of antibody to the immunogen are selected for hybridoma production.

Spleens are removed from immune mice and single-cell suspension is prepared (Harlow et al 1988). Cell fusions are performed essentially as described by Kohler et al.. Briefly, P365.3 myeloma cells (ATTC Rockville, Md) or NS-1 myeloma cells are fused with spleen cells using polyethylene glycol as described by Harlow et al (1989). Cells are plated at a density of 2 x 10^5 cells/well in 96-well tissue culture plates. Individual wells are examined for growth and the supernatants of wells with growth are tested for the presence of complex-specific antibodies by ELISA or RIA using the protein-protein complex of columns 1 and 4 of Table 2 as a target protein. Cells in positive wells are expanded and subcloned to establish and confirm monoclonality.

Clones with the desired specificities are expanded and grown as ascites in mice or in a hollow fiber system to produce sufficient quantities of antibodies for characterization and assay development. Antibodies are tested for binding to bait polypeptide of column 1 of Table 2 alone or to prey polypeptide of column 4 of Table 2 alone, to determine which are specific for the protein-protein complex of columns 1 and 4 of Table 2 as opposed to those that bind to the individual proteins.

Monoclonal antibodies against each of the complexes set forth in columns 1 and 4 of Table 2 are prepared in a similar manner by mixing specified proteins together, immunizing an animal, fusing spleen cells with myeloma cells and isolating clones which produce antibodies specific for the protein complex, but not for individual proteins.

Example 7: Modulating compounds identification

Each specific protein-protein complex of columns 1 and 4 of Table 2 may be used to screen for modulating compounds.

One appropriate construction for this modulating compound screening may be:

- bait polynucleotide inserted in pB27 or pB28;
- prey polynucleotide inserted in pP6;
- transformation of these two vectors in a permeable yeast cell;
- growth of the transformed yeast cell on a medium containing compound to be tested,
- and observation of the growth of the yeast cells.

Example 8 : ZNF8 (hgx554)

GI:17482320

The predicted ZNF8 protein (575 aa) contains 7 zinc finger domains (Lania *et al.*, 1990).

A recent paper has shown mouse ZNF8 (mZNF8) as interacting with the smad1 protein. In addition, mZNF8 was shown to be involved in the TGFβ-BMP pathway (Jiao *et al.*, 2002).

Nucleic acid sequence :

ATGTGT GTGATGTTTC AGGAACCAGT GACCTTCCGG GATGTGGCTG TGGACTTTAC
CCAGGAGGAA TGGGGGCAGC TGGACCCTAC CCAGAGGATC CTCTACCGTG
ACGTGATGCT GGAGACCTTT GGTCACCTGC TCTCCATAGG TCCTGAGCTT
CCGAAGCCTG AAGTCATCTC CCAGCTGGAG CAAGGGACCG AGCTATGGGT
GGCTGAGAGA GGAACCACCC AGGGCTGCCA TCCAGCCTGG GAGCCTCGAT
CTGAAAGCCA AGCATCACGC AAGGAAGAGG GCCTGCCTGA AGAGGAGCCA
TCCCATGTCA CGGGAAGGGA AGGATTCCCG ACAGATGCTC CTTATCCCAC
CACGTTAGGG AAAGACAGGG AGTGTCAGAG CCAGAGTCTG GCACTCAAGG
AGCAGAATAA CTTGAAGCAG TTGGAATTTG GCCTCAAGGA AGCACCAGTT
CAAGATCAAG GCTACAAAAC TCTCAGACTC AGGGAAAACT GCGTCCTGAG
TTCAAGCCCA AATCCATTCC CAGAGATCTC TAGAGGGGAG TATTTGTATA
CTTACGACTC ACAGATTACA GACTCAGAAC ATAACTCCAG CTTAGTCAGT
CAGCAGACAG GCTCCCCAGG AAAACAGCCC GGTGAAAACA GTGACTGTCA
CAGAGATTCC AGTCAGGCCA TTCCAATTAC GGAACTCACA AAAAGCCAGG
TGCAGGACAA ACCCTACAAA TGTACTGACT GTGGGAAGTC GTTTAACCAT
AACGCACACC TCACCGTGCA CAAGAGGATT CATACGGGAG AAAGACCTTA
TATGTGCAAG GAGTGTGGGA AAGCCTTCAG CCAGAACTCC TCCCTCGTCC

AGCATGAGCG CATCCACACT GGAGACAAGC CCTACAAGTG TGCCGAATGT
GGGAAGTCTT TCTGCCATAG TACACACCTT ACCGTCCATC GGAGGATTCA
CACTGGGGAG AAGCCCTATG AGTGTCAGGA CTGTGGGAGG GCCTTCAACC
AGAACTCCTC CCTGGGGCGG CACAAGAGGA CACACACTGG GGAGAAGCCA
TACACCTGCA GTGTGTGTGG GAAATCCTTC TCTCGGACCA CTTGCCTTTT
CCTGCACCTG AGAACTCACA CCGAGGAGAG GCCCTACGAG TGTAACCACT
GCGGGAAGGG CTTCAGGCAC AGCTCATCCC TGGCCCAGCA CCAGCGGAAG
CACGCGGGGG AGAAGCCCTT TGAGTGCCGC CAGAGGCTGA TCTTTGAGCA
GACGCCAGCT CTCACAAAGC ATGAATGGAC AGAAGCCCTG GGCTGTGACC
CACCTTTGAG TCAAGATGAG AGGACTCACC GAAGCGACAG ACCCTTCAAA
TGTAATCAGT GTGGGAAGTG TTTCATTCAG AGCTCTCACC TCATCCGGCA
CCAGATAACT CACACCAGAG AGGAGCAGCC CCATGGGCGA AGCCGGCGGC
GTGAACAATC CTCGAGCAGG AACTCACACC TGGTTCAGCA TCAACACCCG
AACTCCAGAA AGAGCTCTGC AGGCGGAGCA AAGGCAGGGC AGCCGGAAAG
CAGAGCCCTG GCTTTGTTTG ACATCCAAAA AATCATGCAA GAGAAAAACC
CTGTGCACGT TATTGGGGTG GAAGAGCCTT CTGTGGGTGC TTCCATGTTA
TTTGACATCA GAGAATCCAC ATAG (SEQ ID NO.113)

Protein sequence

MDPEDEGVAGVMSVGPPAARLQEPVTFRDVAVDFTQEEWGQLDPTQRILYRDVMLETFGH
LLSIGPELPKPEVISQLEQGTELWVAERGTTQGCHPAWEPRSESQASRKEEGLPEEEPSHV
TGREGFPTDAPYPTTLGKDRECQSQSLALKEQNNLKQLEFGLKEAPVQDQGYKTLRLREN
CVLSSSPNPFPEISRGEYLYTYDSQITDSEHNSSLVSQQTGSPGKQPGENSDCHRDSSQAI
PITELTKSQVQDKPYKCTDCGKSFNHNAHLTVHKRIHTGERPYMCKECGKAFSQNSSLVQH
ERIHTGDKPYKCAECGKSFCHSTHLTVHRRIHTGEKPYECQDCGRAFNQNSSLGRHKRTHT
GEKPYTCSVCGKSFSRTTCLFLHLRTHTEERPYECNHCGKGFRHSSSLAQHQRKHAGEKP
FECRQRLIFEQTPALTKHEWTEALGCDPPLSQDERTHRSDRPFKCNQCGKCFIQSSHLIRH
QITHTREEQPHGRSRRREQSSSRNSHLVQHQHPNSRKSSAGGAKAGQPESRALALFDIQKI
MQEKNPVHVIGVEEPSVGASMLFDIREST (SEQ ID No.114)

I. ZNF8 interacts with several members of the BMP and TGFβ pathways

By two-hybrid screening in yeast (Placenta library) it was shown that ZNF8 interacts with several members of the BMP pathway:

Smad1-ZNF8

SID : Nucleic sequence, SEQ ID No.27 and Proteic sequence, SEQ ID No. 65

SID : Nucleic sequence, SEQ ID No.31 and Proteic sequence, SEQ ID No. 69

Smad5-ZNF8

SID : Nucleic sequence, SEQ ID No.42 and Proteic sequence, SEQ ID No. 80

Smad9-ZNF8

SID : Nucleic sequence, SEQ ID No.45 and Proteic sequence, SEQ ID No. 83

In addition, ZNF8 was also found interacting with smad proteins using other libraries

Smad1-ZNF8

SID : Nucleic sequence, SEQ ID No.28, 29, 30 and Proteic sequence, SEQ ID No. 66, 67, 68

SID : Nucleic sequence, SEQ ID No.32, 33, 34 and Proteic sequence, SEQ ID No. 70, 71, 72

Smad4-ZNF8

SID : Nucleic sequence, SEQ ID No.38 and Proteic sequence, SEQ ID No. 76

Rebound screening experiments using ZNF8 as bait (nt 732-1301) on Placenta library allowed us to confirm the Smad1-ZNF8 and Smad9-ZNF8 interactions

ZNF8-Smad1

ZNF8-Smad9

In summary, Yeast-two-hybrid screens show that amino-acids 22-268 from Smad1 (SEQ ID No.14) interact with amino-acids 354-433 from ZNF8 (SEQ ID No.114) (see . 11A). Amino-acids 1-152 from Smad4 (SEQ ID No.17) interact with amino-acids 172-441 from ZNF8 (see fig. 11B). Amino-acids 1-268 from Smad5 (SEQ ID No.19) interact with amino-acids 276-437 from ZNF8 (see fig. 11C). Finally, amino-acids 1-233 from Smad9 (SEQ ID No.20) interact with amino-acids 208-1209 from ZNF8 (see fig. 11D).

Interestingly, the full-length ZNF8 protein used as bait behaved as autoactivator. This finding as well as the presence of 7 zinc binding domains led us to hypothesise that ZNF8 could be a transcription factor.

II. ZNF8 is an essential player in the TGFβ and BMP pathways

In order to validate ZNF8's involvement in the TGFβ/BMP pathways, ZNF8 cellular knock-down experiments were performed using chemically synthesised siRNA duplexes (see Materials & Methods for siRNA sequences and protocols). While transiently co-transfecting HepG2 cells using the $p(GTCT)_8$-MLP-Luc reporter and ZNF8-targeting siRNA duplex, a specific dose-dependant repression of the TGFβ-dependant reporter activity was observed (see Fig. 12A) demonstrating a function for ZNF8 in the response to the TGFβ pathway. The repressive effect of ZNF8-targeting siRNA duplex was observed at low concentration of siRNA duplex (4nM) and was enhanced at higher concentrations (40nM). While transiently co-transfecting HepG2 cells using the $p(GC)_{12}$-MLP-Luc reporter and ZNF8-targeting siRNA duplex, a specific dose-dependant repression of the BMP-dependant reporter activity was observed (see Fig. 12B) demonstrating a function for ZNF8 on the response to the BMP

pathway on a minimal BMP responsive element. Similar results were obtained using either BMP6 instead of BMP7 (see Fig. 12C). Modulation of these TGFβ/BMP luciferase reporter activities using ZNF8 cellular knock-down suggest an implication of this putative transcription factor in the regulation of these two pathways.

In order to further elucidate its role on the expression of genes naturally controlled by TGFβ and/or BMPs in cells, a similar siRNA-mediated knock-down experiments followed by quantitative PCR (Q-PCR) analysis of TGFβ or BMP-dependant markers was performed. PAI-1 was a well-known target of TGFβ and was strongly induced by TGFβ in many cell types (*Keeton et al.*, 1991). Osteoblastic differentiation was characterized by expression of alkaline phosphatase as an early pre-osteoblastic marker and alkaline phosphatase transcription is directly controled by BMP signals (Wagner EF and Karsenty G, 2001). Modulation of AP1/jun expression by TGFβ is a cell-type specific phenomenon as TGFβ activates c-jun expression only in epithelial cells, whereas it induces junB in mesenchymal cells. JunB is also an immediate early gene induced by BMP-2 (Mauviel *et al.*, 1996; Chalaux *et al.*, 1998).

Endogenous levels of alkaline phosphatase and junB mRNA were specifically and dose-dependently decreased following transient transfection of ZNF8-targeting siRNA duplex in HepG2 cells treated with BMP7 (see Fig. 13A, 13B & 13C respectively). As expected, endogenous PAI-1 mRNA levels were not affected following the same transfection experiments induced by BMP7 (see Fig. 14A). Expression levels of various controls were not affected at all following the same ZNF8-targeting siRNA duplex transfection: hGUS (human beta-glucuronidase, Oshima *et al.*, 1987, see Fig. 14B) HPRT (hypoxanthine-guanine phosphoribosyltransferase, Patel *et al.*, 1986, data not shown), GAPDH (glyceraldehyde-3-phosphate dehydrogenase, Allen *et al.*, 1987, data not shown) and 18S ribosomal RNA (Schmittgen *et al.*, 2000, data not shown).

Example 9 : LAPTm5 (hgx596)

GI:1255239

Using subtractive hybridization, Adra *et al.* (1996) cloned the cDNA coding for a gene preferentially expressed in adult hematopoietic tissues. The predicted protein (262 amino-acids) contained 5 highly hydrophobic transmembrane domains. Immuno-cytological and cell fractionation studies with a specific antibody revealed a protein localizing in lysosomes. In addition, the gene, named LAPTm5, was found to interact with ubiquitin. Recently, a new rat gene which exhibits 80% of identity with LAPTm5, called GCD-10, was identified as activated in response to neuronal apoptosis (Origasa *et al.*, 2001). LAPTm5 has also been found to be an immediate-early gene induced by retinoic acid during granulocytic differentiation in murine retinoic acid-inducible MPRO promyelocyte cell line (Scott *et al.*, 1996).

Finally, LAPTm5 was shown to be up-regulated in the Sjögren's syndrome which is a chronic autoimmune disease (Azuma *et al.*, 2002) and to be co-expressed with activated macrophage genes in rheumatoid arthritis (Walker *et al.*, 2002). Despite being structurally highly related to a family of lysosomal transporter proteins shown to regulate cellular multidrug resistance (Cabrita *et al.*, 1999; Hogue *et al.*, 1999), no function has been attributed to this gene, yet.

Nucleic acid sequence

ATGGACCCCC GCTTGTCCAC TGTCCGCCAG ACCTGCTGCT GCTTCAATGT
CCGCATCGCA ACCACCGCCC TGGCCATCTA CCATGTGATC ATGAGCGTCT
TGTTGTTCAT CGAGCACTCA GTAGAGGTGG CCCATGGCAA GGCGTCCTGC
AAGCTCTCCC AGATGGGCTA CCTCAGGATC GCTGACCTGA TCTCCAGCTT
CCTGCTCATC ACCATGCTCT TCATCATCAG CCTGAGCCTA CTGATCGGCG
TAGTCAAGAA CCGGGAGAAG TACCTGCTGC CCTTCCTGTC CCTGCAAATC
ATGGACTATC TCCTGTGCCT GCTCACCCTG CTGGGCTCCT ACATTGAGCT
GCCCGCCTAC CTCAAGTTGG CCTCCCGGAG CCGTGCTAGC TCCTCCAGTT
CCCCCTGATG ACGCTGCAGC TGCTGGACTT CTGCCTGAGC ATCCTGACCC
TCTGCAGCTC CTACATGGAA GTGCCCACCT ATCTCAACTT CAAGTCCATG
AACCACATGA ATTACCTCCC CAGCCAGGAG GATATGCCTC ATAACCAGTT
CATCAAGATG ATGATCATCT TTTCCATCGC CTTCATCACT GTCCTTATCT
TCAAGGTCTA CATGTTCAAG TGCGTGTGGC GGTGCTACAG ATTGATCAAG
TGCATGAACT CGGTGGAGGA GAAGAGAAAC TCCAAGATGC TCCAGAAGGT
GGTCCTGCCG TCCTACGAGG AAGCCCTGTC TTTGCCATCG AAGACCCCAG
AGGGGGGCCC AGCACCACCC CCATACTCAG AGGTGTGA (SEQ ID No.115)

Protein sequence

MDPRLSTVRQTCCCFNVRIATTALAIYHVIMSVLLFIEHSVEVAHGKASCKLSQMGYLRIADLI
SSFLLITMLFIISLSLLIGVVKNREKYLLPFLSLQIMDYLLCLLTLLGSYIELPAYLKLASRSRASS
SKFPLMTLQLLDFCLSILTLCSSYMEVPTYLNFKSMNHMNYLPSQEDMPHNQFIKMMIIFSIA
FITVLIFKVYMFKCVWRCYRLIKCMNSVEEKRNSKMLQKVVLPSYEEALSLPSKTPEGGPAP
PPYSEV (SEQ ID No.116)

I. LAPTm5 interacts with Smurf2, a protein involved in the TGFβ pathway

We showed by two-hybrid screening in yeast that LAPTm5 interacts with Smurf2, a E3 ubiquitin ligase known to regulate the protein level of Smad1, 2, 7, SnoN and the TGFβ-activated type I receptor (TβRI).

Smurf2-LAPTm5

SID : Nucleic sequence, SEQ ID No.47, 48, 49 and Proteic sequence, SEQ ID No. 85, 86, 87

Rebound screening experiments using LAPTm5 as bait (nt 654-786) on placenta library allowed us to confirm the Smurf2-LAPTm5 interaction:

LAPTm5-Smurf2

Thus, yeast-two-hybrid screens showed that amino-acids 234-335 from Smurf2 (SEQ ID No.22) interact with amino-acids 251-262 from LAPTm5 (SEQ ID No.116) (see Fig.15).

II. LAPTm5 modulates the TGFβ pathway

The two-hybrid screening results led the involvement of the LAPTm5 protein in the TGFβ pathway. To demonstrate a functional effect in mammalian cells, the LAPTm5 c-DNA was cloned into the pV3 vector and used in our TGFβ reporter assay (see Materials & Methods). Over-expression of LAPTm5 (2 and 10 ng of pV3-LAPTm5) results in a dose dependant 2-fold decrease of TGFβ signaling in HepG2 (Figure 16A). In addition, a similar LAPTm5 over-expression in HEK293 cells (0.5, 2, 10 and 50 ng) results also in a 2-fold decrease of TGFβ signaling (Figure 16B). This LAPTm5 effect was not observed when the BMP signaling and the pGL3-control were tested thus showing a reproducible and specific effect of LAPTm5 (Fig. 16A & B, right panel for both figures).

Next investigated was the endogenous level of LAPTm5 mRNA using Q-PCR. LAPTm5 mRNA was barely detectable in HepG2, HeLa and WI38 cells. In contrast, a strong amount of LAPTm5 mRNA was observed in hematopoietic cells such as CEM, CEMC7, K562 and Jurkat cells (Figure 17A).

Next investigated was the effect of TGFβ on the endogenous level of LAPTm5 mRNA in HepG2 cells. After TGFβ induction for 18 H, a 50-fold induction of LAPTm5 mRNA (Figure 17B) was observed. This induction was TGFβ-specific since no effect was observed using several members of BMP, such as BMP2, 4 and 7 (data not shown).

To confirm LAPTm5 induction by TGFβ in HepG2 cells, a TβRI-targeting siRNA duplex previously shown to dramatically reduce T RI mRNA levels in HepG2 cells (data not shown) and to inhibit the TGFβ pathway (see Fig 13A and Materials & Methods) was transiently transfected into HepG2 cells. Following quantitative PCR analysis of total RNA, the TGFβ induction of LAPTm5 mRNA was totally abolished thus confirming the regulation of LAPTm5 mRNA expression by TGFβ (Figure 17B, right panel).

In order to demonstrate LAPTm5's involvement in the TGFβ/BMP pathways in a functional cellular assay , LAPTm5 cellular knock-down experiments were performed using chemically synthesised siRNA duplexes (see Materials & Methods for siRNA sequences and protocols). While transiently co-transfecting HepG2 cells using the p(GTCT)8-MLP-Luc reporter and LAPTm5-targeting siRNA duplex, a specific dose-dependant activation of the TGFβ-dependant reporter activity was observed (see Fig. 18 A) demonstrating a function for LAPTm5 in the response to the TGFβ pathway. The activating effect of LAPTm5-targeting siRNA duplex was observed at low concentration of siRNA duplex (4nM) and was enhanced

at higher concentrations (40nM). While transiently co-transfecting HepG2 cells using the p(GC)12-MLP-Luc reporter and LAPTm5-targeting siRNA duplex, a specific, dose-dependant and BMP-dependant activation of the BMP-dependant reporter activity was observed (see Fig. 18 B) demonstrating a function for LAPTm5 in the response to the BMP pathway.

In order to further elucidate its role on the expression of genes naturally controlled by TGFβ and/or BMPs in cells, a similar siRNA-mediated knock-down experiments followed by quantitative PCR (Q-PCR) analysis of TGFβ or BMP-dependant markers were performed. Endogenous levels of PAI-1, junB and alkaline phosphatase mRNA were specifically and dose-dependently increased following transient transfection of LAPTm5-targeting siRNA duplex in HepG2 cells treated with either TGFβ (PAI-1 and junB, see Fig. 19 A & B) or BMP7 (alkaline phosphatase, see fig 19 C). Expression levels of various controls were not significantly affected following the same LAPTm5-targeting siRNA duplex transfection: hGUS (see fig 19 D) HPRT, GAPDH and 18S (data not shown).

The inhibition effect of LAPTm5 on the TGFβ pathway as well as the up-regulation of the LAPTm5 mRNA level by TGFβ led us to conclude that LAPTm5 is involved in the negative feedback of the TGFβ signalling. It has been suggested by Kavsak *and coll.* (Kavsak *et al.* 2000) that Smurf2 could address the TGFβ receptors and smad7 to the lysosome for degradation. Thus, by interacting with smurf2, a specific E3 ubiquitin ligase known to be involved in the degradation of the TGFβ receptors, Smad1, Smad2, Smad3, Smad7 and SnoN, LAPTm5 could be a smurf2 receptor in the lysosomal membrane and could address some TGFβ signaling members to the lysosomal compartment to induce their degradation.

Example 10 : RNF11 (hgx555)

GI:7657519

Seki *et al.* (1999) identified a new member of the RING finger family, named RNF11 (154 amino acids). Recently, a differential display analysis of gene expression using NIH 3T3 cells expressing the RET-MEN2A or RET-MEN2B mutant proteins was performed. These germ-line point mutations of the RET gene are responsible for multiple endocrine neoplasia (MEN) type 2A and 2B that develop medullary thyroid carcinoma and pheochromocytoma. It has been shown that RNF11 was up-regulated in these mutant cells (Watanabe *et al.*, 2002). In addition, GNDF was found to up-regulate RNF11 levels (Watanabe *et al.*, 2002). However, no function for RNF11 has been attributed yet.

Nucleic acid sequence

ATGGGGAACT GCCTCAAATC CCCCACCTCG GATGACATCT CCCTGCTTCA
CGAGTCTCAG TCCGACCGGG CTAGCTTTGG CGAGGGGACG GAGCCGGATC
AGGAGCCGCC GCCGCCATAT CAGGAACAAG TTCCAGTTCC AGTCTACCAC
CCAACACCTA GCCAGACTCG GCTAGCAACT CAGCTGACTG AAGAGGAACA

AATTAGGATA GCTCAAAGAA TAGGTCTTAT ACAACATCTG CCTAAAGGAG
TTTATGACCC TGGAAGAGAT GGATCAGAAA AAAAGATCCG GGAGTGTGTG
ATCTGTATGA TGGACTTTGT TTATGGGGAC CCAATTCGAT TTCTGCCGTG
CATGCACATC TATCACCTGG ACTGTATAGA TGACTGGTTG ATGAGATCCT
TCACGTGCCC CTCCTGCATG GAGCCAGTTG ATGCAGCACT GCTTTCATCC
TATGAGACTA ATTGA (SEQ ID No.117)

Protein sequence
MGNCLKSPTSDDISLLHESQSDRASFGEGTEPDQEPPPPYQEQVPVPVYHPTPSQTRLAT
QLTEEEQIRIAQRIGLIQHLPKGVYDPGRDGSEKKIRECVICMMDFVYGDPIRFLPCMHIYHL
DCIDDWLMRSFTCPSCMEPVDAALLSSYETN (SEQ ID No.118)

RNF11 interacts with SARA and Smurf2, proteins involved in the TGFβ pathway

By two-hybrid screening in yeast it was shown that RNF11 interacts with SARA, the "Smad anchoring for Receptor Activation", and Smurf2, a E3 ubiquitin ligase known to regulate the protein level of Smad1, 2, 7, SnoN and TβRI.

Smurf2-RNF11

SID : Nucleic sequence, SEQ ID No.50, 51, 52, 53 and Proteic sequence, SEQ ID No. 88, 89, 90, 91

SARA-RNF11

SID : Nucleic sequence, SEQ ID No.54, 55 and Proteic sequence, SEQ ID No. 92, 93

Rebound screening experiments using truncated RNF11 as bait (hgx555v1: nt 93-462) and Full-length RNF11 as bait (hgx555v2) on placenta library allowed us to confirm the Smurf2-RNF11 interaction and to find a new interaction: Smurf1-RNF11:

RNF11-Smurf2

RNF11-Smurf1

Thus, yeast-two-hybrid screens showed that amino-acids 239-335 from Smurf2 (SEQ ID No.22) (aa 239-335) interact with amino-acids 31-84 from RNF11 (SEQ ID No.118) (see Fig. 20 A). Amino-acids 665-1323 from SARA (SEQ ID No.23) interact with amino-acids 61-154 from RNF11 (see Fig. 20B) and amino-acids 236-415 from Smurf1 interact with amino-acids 31-154 from RNF11 (see Fig. 20 C).

II. RNF11 regulates the SARA protein level

Since E3 ubiquitin-protein ligase activity is likely to be a general function of the RING finger, the association between SARA and RNF11 thus raised the interesting possibility that RNF11 might function to regulate the protein level of SARA. To test this, SARA was over-expressed in HepG2 cells (300 ng of pCDNA-SARA, a gift from Azzeddine ATFI) in the presence and absence of RNF11 (300 ng of pV3-RNF11) and TGFβ (5 ng/ml for 18H) and examined the SARA protein level using anti-SARA antibody (cf Materials & Methods). Figure 21 shows that

the SARA protein level was increased in the presence of RNF11. However, no effect of TGFβ was observed in these conditions. In conclusion, this experiment showed that RNF11 was likely involved in regulation of the SARA protein level.

Example 11 : KIAA1196 (hgx559)

GI: 18591703

Nagase *et al.* (1999) newly determined the sequences of 100 cDNA clones of unknown human genes, named KIAA1193 to KIAA1292, from two sets of size-fractionated human adult and fetal brain cDNA libraries. Among these unknown human genes, the hypothetical zinc finger protein KIAA1196 was identified. Since this putative protein contains 7 zinc fingers (C2H2 type), it has been suspected that it may function as a transcription factor. Moreover, KIAA1196 contains a leucine zipper motif in the domain that we have discovered interacting with Smad1 and is predicted to be a nuclear protein, reinforcing its potential function as a transcription factor. However, no function for this protein has been attributed, yet.

Nucleic acid sequence

ATGCCGGTGG TCCGTGGTGG ACAGACAGTG CCCGGCCAGG CCCCTCTCTG
CTTTGACCCG GGAAGTCCAG CCAGTGACAA GACAGAAGGG AAGAAAAAGG
GGCGGCCAAA AGCCGAGAAC CAGGCCCTCC GAGACATTCC TCTCTCCCTG
ATGAACGACT GGAAGGATGA GTTCAAGGCA CACTCGAGGG TGAAGTGTCC
AAACTCAGGG TGCTGGCTGG AGTTCCCCAG CATCTACGGG CTCAAGTACC
ATTACCAGCG GTGCCAAGGG GGTGCCATCT CAGATCGCCT GGCCTTCCCC
TGCCCCTTCT GCGAGGCCGC ATTCACCTCT AAGACCCAGC TGGAGAAACA
CCGGATCTGG AACCACATGG ACCGACCCCT GCCTGCCTCC AAGCCTGGGC
CCATCAGCAG GCCGGTCACC ATCAGCCGGC CTGTTGGGGT CAGCAAGCCC
ATCGGAGTGA GCAAACCTGT CACTATTGGC AAACCTGTGG GTGTCAGCAA
ACCCATTGGC ATCAGCAAGC CAGTCTCGGT CGGCAGACCC ATGCCAGTCA
CCAAGGCCAT CCCGGTCACT AGGCCCGTGC CAGTCACCAA ACCTGTCACA
GTCAGCAGGC CCATGCCCGT CACCAAGGCC ATGCCGGTCA CCAAACCCAT
CACAGTCACC AAGTCTGTGC CGGTCACCAA ACCCGTACCT GTCACCAAAC
CCATTACGGT AACAAAGCTT GTGACAGTTA CGAAACCCGT GCCGGTCACC
AAGCCAGTGA CAGTCAGCAG GCCCATTGTG GTCAGCAAGC CGGTGACAGT
CAGCAGGCCC ATTGCTATCA GCAGACACAC ACCGCCCTGC AAAATGGTGC
TGCTGACCAG GTCGGAGAAC AAAGCACCTC GTGCCACAGG GAGGAACAGT
GGTAAGAAAA GGGCTGCGGA CAGCCTGGAC ACCTGCCCAA TTCCACCCAA
GCAGGCCAGG CCAGAGAATG GGGAGTACGG CCCCTCCTCC ATGGGCCAGA
GCTCGGCCTT CCAGCTGAGT GCAGACACCA GCAGTGGCTC CTTGTCGCCA
GGCAGCAGGC CGTCAGGGGG CATGGAGGCA CTGAAGGCTG CAGGCCCTGC

GTCCCCGCCT GAGGAGGACC CGGAGCGCAC AAAGCACAGA AGGAAACAGA
AAACACCCAA AAAGTTTACA GGGGAGCAGC CATCCATCTC AGGGACCTTT
GGGCTCAAAG GCCTGGTCAA AGCTGAGGAC AAGGCCCGAG TTCACCGCTC
CAAGAAGCAG GAGGGGCCAG GCCCTGAGGA CGCCCGGAAG AAGGTGCCAG
CTGCCCCCAT CACTGTCAGC AAGGAGGCAC CGGCCCCTGT GGCCCACCCA
GCTCCAGGTG GCCCTGAAGA GCAGTGGCAG AGGGCCATCC ATGAGCGCGG
GGAAGCCGTC TGCCCCACCT GCAACGTGGT CACCCGGAAG ACTCTCGTGG
GGCTTAAGAA GCACATGGAG GTGTGTCAGA AGCTTCAGGA TGCACTCAAG
TGCCAGCACT GCCGGAAGCA GTTCAAGTCC AAAGCCGGCC TCAACTACCA
CACTATGGCC GAGCACAGTG CCAAGCCCTC TGACGCCGAG GCCTCCGAAG
GGGGCGAGCA GGAGGAGCGC GAGAGGCTGC GCAAGGTGCT GAAGCAGATG
GGACGGCTGC GCTGCCCCCA GGAGGGTTGC GGGGCTGCCT TCTCCAGCCT
CATGGGCTAC CAGTACCACC AGCGGCGCTG CGGGAAGCCG CCCTGCGAGG
TGGACAGCCC CTCCTTCCCC TGCACCCACT GTGGCAAGAC GTACCGATCC
AAGGCTGGCC ACGACTACCA CGTGCGCTCG GAGCACACGG CCCCCCCCCC
TGAGGAGCCC ACAGACAAGT CCCCTGAGGC TGAGGACCCG CTGGGTGTGG
AGCGGACCCC AAGCGGGCGT GTCCGCCGCA CGTCGGCCCA GGTGGCGGTG
TTCCACCTGC AGGAGATAGC GGAGGACGAG CTGGCCCGCG ACTGGACCAA
GCGGCGCATG AAGGATGACC TTGTGCCCGA GACCTCACAG CTCAACTACA
CTCGACCAGG GCTCCCCACG CTGAACCCCC AGCTGCTAGA GGCATGGAAG
AATGAAGTGA AGGAGAAAGG CCACGTCAAC TGTCCCAACG ACTGCTGTGA
AGCCATCTAC TCCAGCGTGT CCGGACTCAA GGCTCATCTC GCCAGCTGCA
GTAAGGGGGC CCACCTGGCA GGGAAGTACC GCTGTCTGCT GTGTCCGAAG
GAGTTCAGTT CTGAGAGTGG CGTCAAATAC CACATCCTGA AGACCCACGC
AGAGAACTGG TTCCGAACAT CAGCAGACCC ACCTCCCAAA CACAGGAGCC
AGGACTCATT GGTGCCCAAG AAGGAAAAGA AGAAAAATCT GGCAGGTGGA
AAGAAGCGGG GCCGAAAGCC CAAGGAGCGG ACCCCAGAGG AGCCTGTGGC
CAAGCTGCCC CCGCGCCGGG ACGACTGGCC TCCAGGATGC AGAGACAAGG
GGGCCCGGGG CTCCACCGGC CGGAAGGTGG GAGTCAGCAA GGCGCCTGAA
AAGTGA (SEQ ID No.118)

Protein sequence

MPVVRGGQTVPGQAPLCFDPGSPASDKTEGKKKGRPKAENQALRDIPLSLMNDWKDEFKA
HSRVKCPNSGCWLEFPSIYGLKYHYQRCQGGAISDRLAFPCPFCEAAFTSKTQLEKHRIWN
HMDRPLPASKPGPISRPVTISRPVGVSKPIGVSKPVTIGKPVGVSKPIGISKPVSVGRPMPVT
KAIPVTRPVPVTKPVTVSRPMPVTKAMPVTKPITVTKSVPVTKPVPVTKPITVTKLVTVTKPVP
VTKPVTVSRPIVVSKPVTVSRPIAISRHTPPCKMVLLTRSENKAPRATGRNSGKKRAADSLD
TCPIPPKQARPENGEYGPSSMGQSSAFQLSADTSSGSLSPGSRPSGGMEALKAAGPASPP

EEDPERTKHRRKQKTPKKFTGEQPSISGTFGLKGLVKAEDKARVHRSKKQEGPGPEDARK
KVPAAPITVSKEAPAPVAHPAPGGPEEQWQRAIHERGEAVCPTCNVVTRKTLVGLKKHMEV
CQKLQDALKCQHCRKQFKSKAGLNYHTMAEHSAKPSDAEASEGGEQEERERLRKVLKQM
GRLRCPQEGCGAAFSSLMGYQYHQRRCGKPPCEVDSPSFPCTHCGKTYRSKAGHDYHVR
SEHTAPPPEEPTDKSPEAEDPLGVERTPSGRVRRTSAQVAVFHLQEIAEDELARDWTKRR
MKDDLVPETSQLNYTRPGLPTLNPQLLEAWKNEVKEKGHVNCPNDCCEAIYSSVSGLKAHL
ASCSKGAHLAGKYRCLLCPKEFSSESGVKYHILKTHAENWFRTSADPPPKHRSQDSLVPKK
EKKKNLAGGKKRGRKPKERTPEEPVAKLPPRRDDWPPGCRDKGARGSTGRKVGVSKAPE
K (SEQ ID No.119)

KIAA1196 interacts with Smad1 a protein involved in the BMP/TGFβ pathway

By two-hybrid screening in yeast it was shown that KIAA1196 interacts with Smad1, a protein involved in the BMP/TGFβ pathway.

Smad1-KIAA1196

SID : Nucleic sequence, SEQ ID No.35, 36, 37 and Proteic sequence, SEQ ID No. 73, 74, 75

Rebound screening experiments using truncated KIAA1196 as baits (hgx559v1: nt 1455-2322 and hgx559v2: nt 1929-2499) on placenta library allowed us to confirm the Smad1-KIAA1196 interaction:

KIAA1196-Smad1

Thus, yeast-two-hybrid screens showed that amino-acids 242-465 from Smad1 (SEQ ID No.14) interact with amino-acids 643-774 from KIAA1196 (SEQ ID No.14) (see Fig. 22).

II. KIAA1196 modulates the TGFβ signaling

It has been shown that TGFβ binds ALK1 (which induces phosphorylation of Smad1 and 5) and ALK5 (which induces phosphorylation of Smad2 and 3) in transfected COS cells (Ten Dijke *et al.*, 1994). In addition, recent studies have shown that TGFβ regulates the activation state of the endothelium via a fine balance between ALK5 and ALK1 signaling (Goumans *et al.*, 2002). Since KIAA1196 was found interacting with Smad1, it was investigated whether KIAA1196 could be involved in the TGFβ and/or BMP pathways.

In order to validate KIAA1196's involvement in the TGFβ/BMP pathways, KIAA1196 cellular knock-down experiments were performed using chemically synthesised siRNA duplexes (see Materials & Methods for siRNA sequences and protocols). While transiently co-transfecting HepG2 cells using the p(GTCT)8-MLP-Luc reporter and KIAA1196-targeting siRNA duplex, a specific, dose-dependant and TGFβ-dependant repression of the luciferase reporter activity was observed (see Fig. 23 A) demonstrating a function for KIAA1196 in the TGFβ pathway. The repressive effect of KIAA1196-targeting siRNA duplex was already observed at low concentration of siRNA duplex (4nM) and was further enhanced at higher concentrations (40nM). The same transient transfection experiments performed using p(GC)12-MLP-Luc

reporter system and KIAA1196-targeting siRNA-mediated cellular knock-down did not show any impact on the BMP-specific reporter system using BMP6 to activate the pathway (see Fig. 23 B).

SiRNA-mediated KIAA1196 cellular knock-down were also performed in another cell type: HEK293 cells. A specific, dose-dependant and TGFβ-dependant repression of the p(GTCT)8-MLP-Luc reporter activity was also observed (see Fig. 24). The extend of the repression of the TGFβ-dependant reporter activity observed using KIAA1196-targeting siRNA duplex was almost as efficient as the repression obseved using the positive control (TβRI-targeting siRNA duplex). Modulation of the TGFβ luciferase reporter activity using KIAA1196 cellular knock-down demonstrated an essential implication of this putative transcription factor in the regulation of the TGFβ pathway.

In order to further elucidate KIAA1196's role on the expression of genes naturally controlled by TGFβ in mammalian cells, a similar siRNA-mediated knock-down experiments followed by quantitative PCR (Q-PCR) analysis of TGFβ-dependant markers were performed.

Endogenous levels of PAI-1 and junB mRNA were specifically and dose-dependently decreased following transient transfection of KIAA1196-targeting siRNA duplex in HepG2 cells treated with TGFβ (see Fig. 25 A & B). As expected, endogenous alkaline phosphatase mRNA levels were not stimulated following BMP7 treatment and thus were not affected by KIAA1196-targeting siRNA (see Fig. 25 C). Expression levels of various controls were not affected at all following the same KIAA1196-targeting siRNA duplex transfection: hGUS (see Fig. 25 D), HPRT, GAPDH and 18S ribosomal RNA (data not shown).

Endogenous levels of alkaline phosphatase mRNA were barely affected (only a slight decrease) following transient transfection of KIAA1196-targeting siRNA duplex in BMP7 treated HepG2 cells (see Fig. 25 C). Note however that the endogenous levels of TGFβ-induced alkaline phosphatase mRNA is strongly repressed following transient transfection of KIAA1196-targeting siRNA duplex in TGFβ treated HepG2 cells (see Fig. 25 C).

Example 12 : LMO4 (hgx561)

GI: 1914876

LIM-only proteins are transcriptional regulators that function by mediating protein-protein interactions and include the T cell oncogenes LMO1 and LMO2. By screening expression libraries with the LIM interaction domain of NL1/CLIM2/LDB1, Kenny *et al.* (1998) isolated and characterized LMO4, a novel LIM-only gene. The LMO4 gene was further characterized in terms of genomic organization and comparative chromosomal mapping (Tse *et al.*, 1999). LMO4 was found to be a candidate gene associated with prostate cancer progression since LMO4 was down-regulated in prostate cancer (Mousses *et al.*, 2001). In addition, the LMO4 mRNA is over-expressed in human breast cancer cell lines (5 out of 10) and *in situ* hybridization analysis of 177 primary invasive breast carcinomas revealed over-

expression of LMO4 in 56% of specimens (Visvader *et al.*, 2001). Finally, a recent paper describes an interaction between BRCA1 and LMO4. In functional assays, LMO4 was shown to repress BRCA1-mediated transcriptional activation in mammalian cells, suggesting a role for LMO4 as a repressor of BRCA1 activity in breast tissue (Sum *et al.*, 2002). However, no link between LMO4 and the TGFβ/BMP pathways was previously made.

Nucleic acid sequence

ATGGTGAATC CGGGCAGCAG CTCGCAGCCG CCCCCGGTGA CGGCCGGCTC
CCTCTCCTGG AAGCGGTGCG CAGGCTGCGG GGGCAAGATT GCGGACCGCT
TTCTGCTCTA TGCCATGGAC AGCTATTGGC ACAGCCGGTG CCTCAAGTGC
TCCTGCTGCC AGGCGCAGCT GGGCGACATC GGCACGTCCT GTTACACCAA
AAGTGGCATG ATCCTTTGCA GAAATGACTA CATTAGGTTA TTTGGAAATA
GCGGTGCTTG CAGCGCTTGC GGACAGTCGA TTCCTGCGAG TGAACTCGTC
ATGAGGGCGC AAGGCAATGT GTATCATCTT AAGTGTTTTA CATGCTCTAC
CTGCCGGAAT CGCCTGGTCC CGGGAGATCG GTTTCACTAC ATCAATGGCA
GTTTATTTTG TGAACATGAT AGACCTACAG CTCTCATCAA TGGCCATTTG
AATTCACTTC AGAGCAATCC ACTACTGCCA GACCAGAAGG TCTGCTAA (SEQ ID No.120)

Protein sequence

MVNPGSSSQPPPVTAGSLSWKRCAGCGGKIADRFLLYAMDSYWHSRCLKCSCCQAQLGD
IGTSCYTKSGMILCRNDYIRLFGNSGACSACGQSIPASELVMRAQGNVYHLKCFTCSTCRN
RLVPGDRFHYINGSLFCEHDRPTALINGHLNSLQSNPLLPDQKVC (SEQ ID No.121)

LMO4 interacts with Smad9 a protein involved in the BMP pathway

By two-hybrid screening in yeast it was shown that LMO4 interacts with Smad9, a protein involved in the BMP pathway.

Smad9-LMO4

SID : Nucleic sequence, SEQ ID No.44 and Proteic sequence, SEQ ID No. 82

SID : Nucleic sequence, SEQ ID No.46 and Proteic sequence, SEQ ID No. 84

Thus, yeast-two-hybrid screens showed that amino-acids 209-430 from Smad9 (SEQ ID No.20) (aa 209-430) interact with amino-acids 7-125 from LMO4 (SEQ ID No.121) (see Fig. 26).

II. LMO4 modulates BMP signaling

In order to assay LMO4's functional involvement in the TGFβ/BMP pathways, LMO4 cellular knock-down experiments were performed using chemically synthesised siRNA duplexes (see Materials & Methods for siRNA sequences and protocols). While transiently co-transfecting HepG2 cells using the $p(GC)_{12}$-MLP-Luc reporter and LMO4-targeting siRNA duplex, a specific, dose-dependant and BMP7-dependant repression of the BMP-dependant reporter activity was observed (see Fig. 27 A) suggesting a general function for LMO4 in the

response to the BMP7 pathway. Almost similar results were obtained in HepG2 cells using BMP6 instead of BMP7, further reinforcing the BMP-dependant effect of LMO4 siRNA in these cells (see Fig. 27 B).The repressive effect of LMO4-targeting siRNA duplex was observed at low concentration of siRNA duplex (4nM) and was further enhanced at higher concentrations (40nM) for both BMP7 and BMP6 (Fig. 27 A & B).

This effect was shown to be specific and restricted to the BMP pathway since LMO4 did not show any effect on the TGFβ signaling either at 4 or 40 nM of siRNA duplex (see Fig. 27 C). Modulation of the BMP-specific luciferase reporter activity using LMO4 cellular knock-down demonstrates the implication of this putative transcription factor in the regulation of the BMP pathway in HepG2 cells.

In order to further elucidate LMO4's role on the expression of genes naturally controlled by TGFβ and/or BMPs in cells, we performed similar siRNA-mediated knock-down experiments followed by quantitative PCR (Q-PCR) analysis of TGFβ or BMP-dependant markers.

Endogenous levels of alkaline phosphatase mRNA (see Fig. 28 A & B) were specifically and dose-dependently decreased following transient transfection of LMO4-targeting siRNA duplex in HepG2 cells treated with BMP7 demonstrating the role played by LMO4 in the BMP pathway. However, endogenous levels of junB were not affected at all following transient transfection of LMO4-targeting siRNA duplex in HepG2 cells treated with BMP7 (Fig. 29 A). As expected, endogenous PAI-1 mRNA levels were not affected following the same transfection experiments induced by BMP7 (see Fig. 29 C). Expression levels of various controls were not affected at all following the same LMO4-targeting siRNA duplex transfection: hGUS see Fig. 29 B) HPRT, GAPDH and 18S ribosomal RNA (data not shown).

Example 13 : PP1ca (hgx591)

GI: 4506002

Protein phosphatase 1 (PP1) is a major eukaryotic protein serine/threonine phosphatase that regulates an enormous variety of cellular functions through the interaction of its catalytic subunit (PP1c) with over fifty different established or putative regulatory subunits (see for review; Cohen, 2002). Most of these target PP1c to specific sub-cellular locations and interact with a small hydrophobic groove on the surface of PP1c through a short conserved binding motif - the RVxF motif - which is often preceded by further basic residues. Recently, Bennett and Alphey (2002) showed that PP1 binds SARA and negatively regulates Dpp signaling in *Drosophila melanogaster*. Using SARA mutant defective for PP1c binding, they demonstrated that the absence of such interaction resulted in increased expression of TGFβ-reporter gene through increased phosphorylation of type I receptor in the absence of TGFβ.

Nucleic acid sequence

ATGTCCGACA GCGAGAAGCT CAACCTGGAC TCGATCATCG GGCGCCTGCT
GGAAGTGCAG GGCTCGCGGC CTGGCAAGAA TGTACAGCTG ACAGAGAACG
AGATCCGCGG TCTGTGCCTG AAATCCCGGG AGATTTTTCT GAGCCAGCCC
ATTCTTCTGG AGCTGGAGGC ACCCCTCAAG ATCTGCGGTG ACATACACGG
CCAGTACTAC GACCTTCTGC GACTATTTGA GTATGGCGGT TTCCCTCCCG
AGAGCAACTA CCTCTTTCTG GGGGACTATG TGGACAGGGG CAAGCAGTCC
TTGGAGACCA TCTGCCTGCT GCTGGCCTAT AAGATCAAGT ACCCCGAGAA
CTTCTTCCTG CTCCGTGGGA ACCACGAGTG TGCCAGCATC AACCGCATCT
ATGGTTTCTA CGATGAGTGC AAGAGACGCT ACAACATCAA ACTGTGGAAA
ACCTTCACTG ACTGCTTCAA CTGCCTGCCC ATCGCGGCCA TAGTGGACGA
AAAGATCTTC TGCTGCCACG GAGGCCTGTC CCCGGACCTG CAGTCTATGG
AGCAGATTCG GCGGATCATG CGGCCCACAG ATGTGCCTGA CCAGGGCCTG
CTGTGTGACC TGCTGTGGTC TGACCCTGAC AAGGACGTGC AGGGCTGGGG
CGAGAACGAC CGTGGCGTCT CTTTTACCTT TGGAGCCGAG GTGGTGGCCA
AGTTCCTCCA CAAGCACGAC TTGGACCTCA TCTGCCGAGC ACACCAGGTG
GTAGAAGACG GCTACGAGTT CTTTGCCAAG CGGCAGCTGG TGACACTTTT
CTCAGCTCCC AACTACTGTG GCGAGTTTGA CAATGCTGGC GCCATGATGA
GTGTGGACGA GACCCTCATG TGCTCTTTCC AGATCCTCAA GCCCGCCGAC
AAGAACAAGG GGAAGTACGG GCAGTTCAGT GGCCTGAACC CTGGAGGCCG
ACCCATCACC CCACCCCGCA ATTCCGCCAA AGCCAAGAAA TAG (SEQ ID NO122)

Protein sequence

MSDSEKLNLDSIIGRLLEVQGSRPGKNVQLTENEIRGLCLKSREIFLSQPILLELEAPLKICGDI
HGQYYDLLRLFEYGGFPPESNYLFLGDYVDRGKQSLETICLLLAYKIKYPENFFLLRGNHEC
ASINRIYGFYDECKRRYNIKLWKTFTDCFNCLPIAAIVDEKIFCCHGGLSPDLQSMEQIRRIMR
PTDVPDQGLLCDLLWSDPDKDVQGWGENDRGVSFTFGAEVVAKFLHKHDLDLICRAHQVV
EDGYEFFAKRQLVTLFSAPNYCGEFDNAGAMMSVDETLMCSFQILKPADKNKGKYGQFSG
LNPGGRPITPPRNSAKAKK (SEQ ID No.123)

PP1ca interacts with SARA, a protein involved in the TGFβ pathway

By two-hybrid screening in yeast it was shown that PP1ca interacts with SARA, a protein involved in the TGFβ pathway.

SARA-PP1ca

SID : Nucleic sequence, SEQ ID No.56, 57,58, 59 and Proteic sequence, SEQ ID No. 94, 95, 96, 97.

In addition, these screens using SARA as bait gave the two additional isoforms of PP1c: PP1cb and PP1cc.

Rebound screening experiments using PP1ca as bait (hgx591v2: nt 1-972) on a placenta library allowed us to confirm the SARA-PP1ca interaction.

PP1ca-SARA

Thus, yeast-two-hybrid screens showed that amino-acids 668-947 from SARA (SEQ ID No.23) interact with amino-acids 29-295 from PP1ca (SEQ ID No.123) (see Fig. 30).

PP1ca is a regulator of the TGFβ signaling

The two-hybrid screening results led the involvement of the catalytic subunit of the serine/threonine phosphatase (PP1c) protein in the TGFβ pathway. Sequence analysis of SARA revealed the presence of the RVxF motif in the C-terminal part of SARA found to interact with PP1c (aa 668-947). To show a functional involvement of PP1ca in the TGFβ pathway, the *PP1ca* c-DNA (nt 1-972) was cloned into the pV3 vector and used in our TGFβ reporter assay (cf Materials & Methods).

Over-expression of PP1ca at several amounts (10, 50 and 200 ng of pV3-PP1ca) results in a 3.5-fold increase and a 6-fold increase of TGFβ signaling in HepG2 and HEK293 cells, respectively (Figure 31 A & B, respectively). This PP1ca effect was not observed when the BMP signaling and the pGL3-control was tested thus showing a reproducible and specific effect of PP1ca on the TGFβ pathway (Figure 31 A & B respectively; right panels).

To confirm this finding, a Baculovirus over-expressing the smad3 protein (as positive control) and the PP1ca protein was generated. This baculovirus expression system has been genetically engineered to allow infection and expression in mammalian cells (see material & methods). Both viruses were used to infect the HepG2 cells for 24 hours with or without TGFβ. First, the over-expression level of our proteins of interest by Q-PCR experiments was checked. In these conditions, the Smad3 and PP1ca mRNA were shown to be over-expressed by a 350-fold and 50-fold, respectively, as compared to the endogenous mRNA level (Figure 32 A). Next, the endogenous PAI-1 and JunB mRNA levels was looked at, which were previously shown to be up-regulated by TGF . In the case of PAI-1 expression, in the absence of TGFβ, a 5-fold induction by Smad3 and a 2.5-fold induction by PP1ca (Figure 32 B, left panel) was observed. In the presence of TGFβ, a 2.5-fold induction by Smad3 but no effect of PP1ca (Figure 32 B, right panel) was observed. Concerning the Jun-B expression, a 7-fold induction by smad3 was observed in the absence of TGFβ (data not shown). However, no effect was observed following PP1ca over-expression. This result suggests that PP1ca is involved in regulation of PAI-1 expression.

Example 14 : HYPA (hgx530)

GI: 3341989

Huntington's disease, with its hallmark choreiform movements and graded loss of striatal neurons, is a dominantly inherited disorder caused by expansion of a CAG repeat in one copy of the *HD* gene. The *HD* mutation elongates an N-terminal glutamine segment in

the huntingtin protein. HYPA, HYPB and HYPC were found to interact with the huntingtin protein (Faber *et al.*, 1998). HYPA is a protein containing a WW domain, known to bind prolin-rich peptides stretches. This protein is the human homolog of the essential pre-mRNA splicing factor PrP40 and is also called FBP11. Modification of mutant huntingtin in target neurons may promote an abnormal interaction with one, or all, huntingtin's WW domain partners, perhaps altering ribonucleo-protein function with toxic consequences (Passani *et al.*, 2000). In addition, HYPA contains a FF domain, with a structure recently determined, which is a 60 amino acid residue phosphopeptide-binding module (Allen *et al.*, 2002). However, no link between HYPA and the TGFβ/BMP pathway was previously made.

Nucleic acid sequence

CTGAGCCCGA CGATGAGGCC GGGGACGGGA GCTGAGCGTG GAGGCCTCAT
GGTGAGTGAA ATGGAGAGCC ATCCTCCCTC GCAGGGTCCT GGGGACGGGG
AGCGGAGATT GTCCGGCTCA AGCCTCTGCT CCGGCTCTTG GGTCTCTGCT
GACGGCTTCC TGAGGAGACG GCCCTCGATG GGGCACCCTG GCATGCATTA
TGCCCCAATG GGAATGCACC CTATGGGTCA GAGAGCGAAT ATGCCTCCTG
TACCTCATGG AATGATGCCG CAGATGATGC CCCCTATGGG AGGGCCACCA
ATGGGACAAA TGCCTGGAAT GATGTCGTCA GTAATGCCTG GAATGATGAT
GTCTCATATG TCTCAGGCTT CCATGCAGCC TGCCTTACCG CCAGGAGTAA
ATAGTATGGA TGTAGCAGCA GGTACAGCAT CTGGTGCAAA ATCAATGTGG
ACTGAACATA AATCACCTGA TGGAAGGACT TACTACTACA ACACTGAAAC
CAAACAGTCT ACCTGGGAGA AACCAGATGA TCTTAAAACA CCTGCTGAGC
AACTCTTATC TAAATGCCCC TGGAAGGAAT ACAAATCAGA TTCTGGAAAG
CCTTACTATT ATAATTCTCA AACAAAAGAA TCTCGCTGGG CCAAACCTAA
AGAACTTGAG GATCTTGAAG GATACCAGAA TACCATTGTT GCTGGAAGTC
TTATTACAAA ATCAAACCTG CATGCAATGA TCAAAGCTGA AGAAAGCAGT
AAGCAAGAAG AGTGCACCAC AACATCAACA GCCCCAGTCC CTACAACAGA
AATTCCGACC ACAATGAGCA CCATGGCTGC TGCCGAAGCA GCAGCTGCTG
TTGTTGCAGC AGCAGCAGCG GCAGCAGCAG CAGCAGCTGC AGCCAATGCT
AATGCTTCCA CTTCTGCTTC TAATACTGTC AGTGGAACTG TTCCAGTTGT
TCCTGAGCCT GAAGTTACTT CCATTGTTGC TACTGTTGTA GATAATGAGA
ATACAGTAAC TATTTCAACT GAGGAACAAG CACAACTTAC TAGTACCCCT
GCTATTCAGG ATCAAAGTGT GGAAGTATCC AGTAATACTG GAGAAGAAAC
ATCTAAGCAA GAAACTGTAG CTGATTTTAC TCCCAAAAAA GAAGAGGAGG
AGAGCCAACC AGCAAAGAAA ACATACACTT GGAATACAAA GGAAGAGGCA
AAGCAAGCTT TTAAAGAATT ATTGAAAGAA AAGCGGGTAC CATCGAATGC
TTCATGGGAG CAGGCTATGA AAATGATTAT TAATGATCCA CGATACAGTG

CTTTGGCAAA CTTAAGTGAA AAAAAGCAAG CCTTTAATGC CTATAAAGTC CAGACAG (SEQ ID No.124)

Protein sequence

LSPTMRPGTGAERGGLMVSEMESHPPSQGPGDGERRLSGSSLCSGSWVSADGFLRRRP SMGHPGMHYAPMGMHPMGQRANMPPVPHGMMPQMMPPMGGPPMGQMPGMMSSVMP GMMMSHMSQASMQPALPPGVNSMDVAAGTASGAKSMWTEHKSPDGRTYYYNTETKQST WEKPDDLKTPAEQLLSKCPWKEYKSDSGKPYYYNSQTKESRWAKPKELEDLEGYQNTIVA GSLITKSNLHAMIKAEESSKQEECTTTSTAPVPTTEIPTTMSTMAAAEAAAAVVAAAAAAAAA AAAANANASTSASNTVSGTVPVVPEPEVTSIVATVVDNENTVTISTEEQAQLTSTPAIQDQSV EVSSNTGEETSKQETVADFTPKKEEEESQPAKKTYTWNTKEEAKQAFKELLKEKRVPSNAS WEQAMKMIINDPRYSALANLSEKKQAFNAYKVQT (SEQ ID No.125)

HYPA interacts with Smad4, a protein involved in the TGFβ/ BMP pathway

By two-hybrid screening in yeast it was shown that HYPA interacts with Smad4, a protein involved in the TGFβ/BMP pathway.

Smad4-HYPA

SID : Nucleic sequence, SEQ ID No.39, 40, 41 and Proteic sequence, SEQ ID No. 77, 78, 79.

Thus, yeast-two-hybrid screens showed that amino-acids 251-552 from Smad4 (SEQ ID No.17) interact with amino-acids 276-387 from HYPA (SEQ ID No.123) (see Fig. 33).

HYPA is a regulator of the TGFβ signaling

Since HYPA was found interacting with Smad4, it was investigated whether HYPA could be involved in the TGFβ and/or BMP pathways. In order to assay HYPA's functional involvement in the TGFβ/BMP pathways, HYPA cellular knock-down experiments were performed using chemically synthesised siRNA duplexes (see Materials & Methods for siRNA sequences and protocols). While transiently co-transfecting HepG2 cells using the $p(GC)_8$-MLP-Luc reporter and HYPA-targeting siRNA duplex, a specific dose-dependant repression of the BMP-dependant reporter activity was observed (see Fig. 34 A) demonstrating a function for HYPA in the response to the BMP pathway. Similar results were also obtained using either BMP6 or BMP7 (see Fig. 34 A & B). The repressive effect of HYPA-targeting siRNA duplex was observed at low concentration of siRNA duplex (4nM) and was enhanced at higher concentrations (40nM). While transiently co-transfecting HepG2 cells using the $p(GTCT)_{12}$-MLP-Luc reporter and HYPA-targeting siRNA duplex, no repression of the TGFβ-dependant reporter activity was observed (see Fig. 34 C) demonstrating a restricted function for HYPA in the response to the BMP pathway.

In order to further elucidate HYPA's role on the expression of genes naturally controlled by BMPs in mammalian cells, a similar siRNA-mediated knock-down experiments followed by quantitative PCR (Q-PCR) analysis of BMP-dependant markers was performed.

Endogenous levels of alkaline phosphatase mRNA were specifically and dose-dependently decreased following transient transfection of HYPA-targeting siRNA duplex in HepG2 cells treated with BMP7 (see Fig. 35) demonstrating HYPA's role in the BMP pathway. Expression levels of various controls were not affected at all following the same HYPA-targeting siRNA duplex transfection: hGUS, HPRT, GAPDH and 18S ribosomal RNA (data not shown).

Example 15 : FLJ20037 (hgx594)

GI: 8923041

Lagali *et al.* (2002) have identified a novel human gene, chromosome 6 open reading frame 37 (C6orf37), also named FLJ20037, that is expressed in the retina and maps to human chromosome 6q14, a genomic region that harbors multiple retinal disease loci. Northern blot analysis indicates that this gene is widely expressed, with preferential expression observed in the retina compared to other ocular tissues. The C6orf37 protein shares homology with putative proteins in *R. norvegicus, M. musculus, D. melanogaster* and *C. elegans*, suggesting evolutionary conservation of function. Additional sequence analysis predicts that the C6orf37 gene product is a soluble, globular cytoplasmic protein containing several conserved phosphorylation sites. The N-terminal part of this protein contains some glycine-rich repeats. However, no link between FLJ20037 and the TGFβ/BMP pathway was previously made.

Nucleic acid sequence

ATGGCGGAGG GTGAAGGGTA CTTCGCCATG TCTGAGGACG AGCTGGCCTG
CAGCCCCTAC ATCCCCCTAG GCGGCGACTT CGGCGGCGGC GACTTCGGCG
GCGGCGACTT CGGCGGCGGC GACTTCGGCG GCGGCGACTT CGGCGGTGGC
GGCAGCTTCG GTGGGCATTG CTTGGACTAT TGCGAAAGCC CTACGGCGCA
CTGCAATGTG CTGAACTGGG AGCAAGTGCA GCGGCTGGAC GGCATCCTGA
GTGAGACCAT TCCGATTCAC GGGCGCGGCA ACTTCCCCAC GCTCGAGCTG
CAGCCGAGCC TGATCGTGAA GGTGGTGCGG CGGCGCCTGG CCGAGAAGCG
CATTGGCGTC CGCGACGTGC GCCTCAACGG CTCGGCAGCC AGCCATGTCC
TGCACCAGGA CAGCGGCCTG GGCTACAAGG ACCTGGACCT CATCTTCTGC
GCCGACCTGC GCGGGGAAGG GGAGTTTCAG ACTGTGAAGG ACGTCGTGCT
GGACTGCCTG TTGGACTTCT TACCCGAGGG GGTGAACAAA GAGAAGATCA
CACCACTCAC GCTCAAGGAA GCTTATGTGC AGAAAATGGT TAAAGTGTGC
AATGACTCTG ACCGATGGAG TCTTATATCC CTGTCAAACA ACAGTGGCAA
AAATGTGGAA CTGAAATTTG TGGATTCCCT CCGGAGGCAG TTTGAATTCA

GTGTAGATTC TTTTCAAATC AAATTAGACT CTCTTCTGCT CTTTTATGAA
TGTTCAGAGA ACCCAATGAC TGAGACATTT CACCCCACAA TAATCGGGGA
GAGCGTCTAT GGCGATTTCC AGGAAGCCTT TGATCACCTT TGTAACAAGA
TCATTGCCAC CAGGAACCCA GAGGAAATCC GAGGGGGAGG CCTGCTTAAG
TACTGCAACC TCTTGGTGAG GGGCTTTAGG CCCGCCTCTG ATGAAATCAA
GGCCCTTCAA AGGTACATGT GTTCCAGGTT TTTCATCGAC TTCTCAGACA
TTGGAGAGCA GCAGAGAAAA CTGGAGTCCT ATTTGCAGAA CCACTTTGTG
GGATTGGAAG ACCGCAAGTA TGAGTATCTC ATGACCCTTC ATGGAGTGGT
AAATGAGAGC ACAGTGTGCC TGATGGGACA TGAAAGAAGA CAGACTTTAA
ACCTTATCAC CATGCTGGCT ATCCGGGTGT TAGCTGACCA AAATGTCATT
CCTAATGTGG CTAATGTCAC TTGCTATTAC CAGCCAGCCC CCTATGTAGC
AGATGCCAAC TTTAGCAATT ACTACATTGC ACAGGTTCAG CCAGTATTCA
CGTGCCAGCA ACAGACCTAC TCCACTTGGC TACCCTGCAA TTAA (SEQ ID No.124)

Protein sequence

MAEGEGYFAMSEDELACSPYIPLGGDFGGGDFGGGDFGGGDFGGGDFGGGGSFGGHCL
DYCESPTAHCNVLNWEQVQRLDGILSETIPIHGRGNFPTLELQPSLIVKVVRRRLAEKRIGVR
DVRLNGSAASHVLHQDSGLGYKDLDLIFCADLRGEGEFQTVKDVVLDCLLDFLPEGVNKEKI
TPLTLKEAYVQKMVKVCNDSDRWSLISLSNNSGKNVELKFVDSLRRQFEFSVDSFQIKLDSL
LLFYECSENPMTETFHPTIIGESVYGDFQEAFDHLCNKIIATRNPEEIRGGGLLKYCNLLVRGF
RPASDEIKALQRYMCSRFFIDFSDIGEQQRKLESYLQNHFVGLEDRKYEYLMTLHGVVNEST
VCLMGHERRQTLNLITMLAIRVLADQNVIPNVANVTCYYQPAPYVADANFSNYYIAQVQPVF
TCQQQTYSTWLPCN (SEQ ID No.125)

FLJ20037 interacts with SARA, a protein involved in the TGFβ pathway

By two-hybrid screening in yeast it was shown that FLJ20037 interacts with SARA, a protein involved in the TGFβ pathway.

SARA-FLJ20037

SID : Nucleic sequence, SEQ ID No.60, 61 and Proteic sequence, SEQ ID No. 98, 99.

Thus, yeast-two-hybrid screens showed that amino-acids 665-1323 from SARA (SEQ ID No.23) interact with amino-acids 58-253 from FLJ20037 (SEQ ID No.125) (see Fig. 36).

FLJ20037 modulates the TGFβ signaling

Since FLJ20037 was found as interacting with SARA, it was investigated whether FLJ20037 could be involved in the TGFβ and/or BMP pathways. To test this, baculoviruses over-expressing the smad3 protein (as positive control) and the FLJ20037 protein were generated. Both viruses were used to infect the HepG2 cells during 24 hours, treated or not with TGFβ.

First, the over-expression level of our proteins of interest by Q-PCR experiments was checked. Smad3 and FLJ20037 mRNA were shown to be over-expressed 350-fold and 200-fold, respectively, when compared to their respective endogenous mRNA levels (Figure 37 A). Next, endogenous PAI-1 and JunB mRNA levels were looked at, which were previously shown to be up-regulated by TGFβ. In the case of PAI-1 expression, in the absence of TGFβ, a 5-fold induction by smad3 and a 3.5-fold induction by FLJ20037 were observed (Figure 37 B, left panel). In the presence of TGFβ, we observed a 2.5-fold induction by Smad3 and a 2-fold induction by FLJ20037 (Figure 37 B, right panel). Concerning the Jun-B expression, a 7-fold induction by smad3 was observed in the absence of TGFβ (data not shown). However, no effect on junB expression was observed following FLJ20037 over-expression (data not shown). These results suggested that FLJ20037 was involved in the regulation of PAI-1 expression.

In order to further elucidate FLJ20037's role on the expression of genes naturally controlled by TGFβ in mammalian cells, siRNA-mediated knock-down experiments followed by quantitative PCR (Q-PCR) analysis of TGFβ-dependant markers were performed (see Materials & Methods for siRNA sequences and protocols). Endogenous levels of PAI-1 mRNA were specifically decreased following transient transfection of FLJ20037-targeting siRNA duplex in HepG2 cells treated with TGFβ (see Fig. 38). Expression levels of various controls were not affected at all following the same HYPA-targeting siRNA duplex transfection: hGUS, HPRT, GAPDH and 18S ribosomal RNA (data not shown).

Example 16-PTPN12

GI: 18375651

Protein tyrosine phosphatases (PTPs) and protein tyrosine kinases (PTKs) are involved in the regulation of tyrosine phosphorylation-mediated signaling. Such signaling is critical for the regulation of cell proliferation, differentiation, and neoplastic transformation. Tyrosine-phosphorylated proteins can be specifically dephosphorylated through the action of PTPs, which therefore are likely to have as important a role as PTKs in the control of cellular growth and differentiation. Given that hyperphosphorylation of protein tyrosine residues can cause cell transformation, it is plausible that lack of dephosphorylation resulting from loss of PTP function may also wreak an oncogenic effect. Intracellular PTPs are candidates for tumor suppressor genes. From an adult cDNA library, Takekawa et al. (1992) isolated a cDNA encoding a predicted 88-kD protein and Yang et al. (1993) isolated a virtually identical gene from HeLa cell extracts. The protein was designated protein tyrosine phosphatase G1 (PTPG1) or PTPN12 or PTP-PEST. Cong et al. (2000) showed that PSTPIP1 bridges ABL to the PEST-type PTPs. Several experiments suggested that the PEST-type PTPs negatively regulate ABL activity: ABL was hyperphosphorylated in PTP-PEST-deficient cells; disruption of the ABL-PSTPIP1-PEST-type PTP ternary complex by overexpression of mutants

increased ABL phosphotyrosine content; and PDGF-induced ABL kinase activation was prolonged in PTP-PEST-deficient cells. The authors concluded that dephosphorylation of ABL by PSTPIP1-directed PEST-type PTPs represents a novel mechanism by which ABL activity is regulated. Charest et al. (1995) determined that the mouse PTPN12 gene contains 18 exons spanning about 90 kb of DNA. By fluorescence in situ hybridization (FISH), Takekawa et al. (1994) mapped the PTPN12 gene to 7q11.23. Charest et al. (1995) used FISH to map the mouse Ptpn12 gene to chromosome 5A3 to B, a region with homology of synteny to human chromosome 7q11.23. The potential importance of PTPG1 in tumorigenesis was investigated by Takekawa et al. (1994), who sought abnormalities of the PTPG1 transcript in various human cancer cell lines by use of RT-PCR. In a colorectal carcinoma cell line, DLD-1, they found 3 aberrant transcripts (Sequencing in one demonstrated an A-to-G transition at nucleotide 201, predicting a change of codon 61 from lysine to arginine): a missense point mutation, a 77-bp deletion, and a 173-bp deletion. However, no link between HIPK3 and the TGF /BMP pathway was previously made.

Nucleic acid sequence

ATGGAGCAAG TGGAGATCCT GAGGAAATTC ATCCAGAGGG TCCAGGCCAT
GAAGAGTCCT GACCACAATG GGGAGGACAA CTTCGCCCGG GACTTCATGC
GGTTAAGAAG ATTGTCTACC AAATATAGAA CAGAAAAGAT ATATCCCACA
GCCACTGGAG AAAAAGAAGA AAATGTTAAA AAGAACAGAT ACAAGGACAT
ACTGCCATTT GATCACAGCC GAGTTAAATT GACATTAAAG ACTCCTTCAC
AAGATTCAGA CTATATCAAT GCAAATTTTA TAAAGGGCGT CTATGGGCCA
AAAGCATATG TAGCAACTCA AGGACCTTTA GCAAATACAG TAATAGATTT
TTGGAGGATG ATATGGGAGT ATAATGTTGT GATCATTGTA ATGGCCTGCC
GAGAATTTGA GATGGGAAGG AAAAAATGTG AGCGCTATTG GCCTTTGTAT
GGAGAAGACC CCATAACGTT TGCACCATTT AAAATTTCTT GTGAGGATGA
ACAAGCAAGA ACAGACTACT TCATCAGGAC ACTCTTACTT GAATTTCAAA
ATGAATCTCG TAGGCTGTAT CAGTTTCATT ATGTGAACTG GCCAGACCAT
GATGTTCCTT CATCATTTGA TTCTATTCTG GACATGATAA GCTTAATGAG
GAAATATCAA GAACATGAAG ATGTTCCTAT TTGTATTCAT TGCAGTGCAG
GCTGTGGAAG AACAGGTGCC ATTTGTGCCA TAGATTATAC GTGGAATTTA
CTAAAAGCTG GGAAAATACC AGAGGAATTT AATGTATTTA ATTTAATACA
AGAAATGAGA ACACAAAGGC ATTCTGCAGT ACAAACAAAG GAGCAATATG
AACTTGTTCA TAGAGCTATT GCCCAACTGT TTGAAAAACA GCTACAACTA
TATGAAATTC ATGGAGCTCA GAAAATTGCT GATGGAGTGA ATGAAATTAA
CACTGAAAAC ATGATCAGCT CCATAGAGCC TGAAAAACAA GATTCTCCTC
CTCCAAAACC ACCAAGGACC CGCAGTTGCC TTGTTGAAGG GGATGCTAAA
GAAGAAATAC TGCAGCCACC GGAACCTCAT CCAGTGCCAC CCATCTTGAC

ACCTTCTCCC CCTTCAGCTT TTCCAACAGT CACTACTGTG TGGCAGGACA
ATGATAGATA CCATCCAAAG CCAGTGTTGC ATATGGTTTC ATCAGAACAA
CATTCAGCAG ACCTCAACAG AAACTATAGT AAATCAACAG AACTTCCAGG
GAAAAATGAA TCAACAATTG AACAGATAGA TAAAAAATTG GAACGAAATT
TAAGTTTTGA GATTAAGAAG GTCCCTCTCC AAGAGGGACC AAAAAGTTTT
GATGGGAACA CACTTTTGAA TAGGGGACAT GCAATTAAAA TTAAATCTGC
TTCACCTTGT ATAGCTGATA AAATCTCTAA GCCACAGGAA TTAAGTTCAG
ATCTAAATGT CGGTGATACT TCCCAGAATT CTTGTGTGGA CTGCAGTGTA
ACACAATCAA ACAAAGTTTC AGTTACTCCA CCAGAAGAAT CCCAGAATTC
AGACACACCT CCAAGGCCAG ACCGCTTGCC TCTTGATGAG AAAGGACATG
TAACGTGGTC ATTTCATGGA CCTGAAAATG CCATACCCAT ACCTGATTTA
TCTGAAGGCA ATTCCTCAGA TATCAACTAT CAAACTAGGA AAACTGTGAG
TTTAACACCA AGTCCTACAA CACAAGTTGA AACACCTGAT CTTGTGGATC
ATGATAACAC TTCACCACTC TTCAGAACAC CCCTCAGTTT TACTAATCCA
CTTCACTCTG ATGACTCAGA CTCAGATGAA AGAAACTCTG ATGGTGCTGT
GACCCAGAAT AAAACTAATA TTTCAACAGC AAGTGCCACA GTTTCTGCTG
CCACTAGTAC TGAAAGCATT TCTACTAGGA AAGTATTGCC AATGTCCATT
GCTAGACATA ATATAGCAGG AACAACACAT TCAGGTGCTG AAAAAGATGT
TGATGTTAGT GAAGATTCAC CTCCTCCCCT ACCTGAAAGA ACTCCTGAAT
CGTTTGTGTT AGCAAGTGAA CATAATACAC CTGTAAGATC GGAATGGAGT
GAACTTCAAA GTCAGGAACG ATCTGAACAA AAAAAGTCTG AAGGCTTGAT
AACCTCTGAA AATGAGAAAT GTGATCATCC AGCGGGAGGT ATTCACTATG
AAATGTGCAT AGAATGTCCA CCTACTTTCA GTGACAAGAG AGAACAAATA
TCAGAAAATC CAACAGAAGC CACAGATATT GGTTTTGGTA ATCGATGTGG
AAAACCCAAA GGACCAAGAG ATCCACCTTC AGAATGGACA TGA (SEQ ID No.126)

Protein sequence

MEQVEILRKFIQRVQAMKSPDHNGEDNFARDFMRLRRLSTKYRTEKIYPTATGEKEENVKK
NRYKDILPFDHSRVKLTLKTPSQDSDYINANFIKGVYGPKAYVATQGPLANTVIDFWRMIWE
YNVVIIVMACREFEMGRKKCERYWPLYGEDPITFAPFKISCEDEQARTDYFIRTLLLEFQNES
RRLYQFHYVNWPDHDVPSSFDSILDMISLMRKYQEHEDVPICIHCSAGCGRTGAICAIDYTW
NLLKAGKIPEEFNVFNLIQEMRTQRHSAVQTKEQYELVHRAIAQLFEKQLQLYEIHGAQKIAD
GVNEINTENMISSIEPEKQDSPPPKPPRTRSCLVEGDAKEEILQPPEPHPVPPILTPSPPSAF
PTVTTVWQDNDRYHPKPVLHMVSSEQHSADLNRNYSKSTELPGKNESTIEQIDKKLERNLS
FEIKKVPLQEGPKSFDGNTLLNRGHAIKIKSASPCIADKISKPQELSSDLNVGDTSQNSCVDC
SVTQSNKVSVTPPEESQNSDTPPRPDRLPLDEKGHVTWSFHGPENAIPIPDLSEGNSSDINY
QTRKTVSLTPSPTTQVETPDLVDHDNTSPLFRTPLSFTNPLHSDDSDSDERNSDGAVTQNK
TNISTASATVSAATSTESISTRKVLPMSIARHNIAGTTHSGAEKDVDVSEDSPPPLPERTPES

FVLASEHNTPVRSEWSELQSQERSEQKKSEGLITSENEKCDHPAGGIHYEMCIECPPTFSD KREQISENPTEATDIGFGNRCGKPKGPRDPPSEWT (SEQ ID No.127)

PTPN12 interacts with Smad5, a protein involved in the BMP pathway

By two-hybrid screening in yeast it was shown that PTPN12 interacts with Smad5, a protein involved in the BMP pathway.

Smad5-PTPN12

SID : Nucleic sequence, SEQ ID No.43 and Proteic sequence, SEQ ID No. 81.

Thus, yeast-two-hybrid screens showed that amino-acids 1-268 from Smad5 (SEQ ID No.19) interact with amino-acids 99-337 from PTPN12 (SEQ ID No.127) (see Fig. 39).

PTPN12 modulates the TGFβ and BMP signaling

Since PTPN12 was found interacting with Smad5, it was investigated whether PTPN12 could be involved in the TGFβ and/or BMP pathways. In order to assay PTPN12's functional involvement in the TGFβ/BMP pathways, PTPN12 cellular knock-down experiments were performed using chemically synthesised siRNA duplexes (see Materials & Methods for siRNA sequences and protocols). While transiently co-transfecting HepG2 cells using the $p(GC)_{12}$-MLP-Luc reporter and PTPN12-targeting siRNA duplex, a specific BMP6-dependant increase in the BMP-dependant reporter activity was observed (see Fig. 40 A) demonstrating a role for PTPN12 in the response to the BMP pathway. The positive effect of PTPN12-targeting siRNA duplex was already observed at low concentration of siRNA duplex (4nM). While transiently co-transfecting HepG2 cells using the $p(GTCT)_8$-MLP-Luc reporter and PTPN12-targeting siRNA duplex, a TGFβ-dependant increase in the reporter activity was also observed (see Fig. 40 B) demonstrating a specific function for PTPN12 on the response to both the TGFβ and BMP pathways. Modulation of the TGF and BMP luciferase reporter activities using PTPN12 cellular knock-down shows its an implication in the regulation of both pathway.

Example 17-HIPK3

GI: 11386208

Recently was identified a 130-kD kinase designated Fas-interacting serine/threonine kinase/homeodomain-interacting protein kinase (FIST/HIPK3) as a novel Fas-interacting protein (Rochat-Steiner *et al.,* 2000). These authors demonstrated that these results suggest that Fas-associated FIST/HIPK3 modulates one of the two major signaling pathways of Fas. Using PCR with degenerate primers based on conserved domains of serine-threonine kinases, Begley et al. (1997) isolated an MDR cell cDNA encoding a 1,215-amino acid protein with a calculated molecular mass of 130 kD. The protein contains sequences identical to the catalytic core of many serine-protein kinases and is 54% similar to the yeast protein kinase YAK1, whose normal role is to restrict growth. The authors therefore designated the protein PKY/HIPK3, for homolog of protein kinase YAK1. The authors stated

that PKY/HIPK3 may be identical to a 170-kD kinase identified in the same cell lines by Sampson et al. (1993), the difference in molecular mass being due to posttranslational modifications. By Northern blot analysis, PKY/HIPK3 was expressed at higher levels in MDR cells than in their nonresistant parental lines; in addition, a 7-kb PKY/HIPK3 transcript was expressed at high levels in heart and skeletal muscle and at lower levels in placenta, pancreas, and brain. Using a yeast 2-hybrid screen, Kim et al. (1998) identified in mouse 3 members of a family of cofactors, which they designated homeodomain-interacting protein kinases (HIPKs), that interact with homeoproteins and show the greatest similarity to the yeast YAK1 protein (43% identity in the catalytic domain). The corepressor activity of HIPKs depends on both its homeodomain interaction domain and a corepressor domain that maps to the N terminus. Kim et al. (1998) presented evidence that HIPKs can act as transcriptional corepressors for NK homeodomain transcriptionfactors. By fluorescence in situ hybridization, Nupponen and Visakorpi (1999) mapped the HIPK3 gene to chromosome 11p13. However, no link between HIPK3 and the TGFβ/BMP pathway was previously made.

Nucleic acid sequence

ATGGCCTCAC AAGTCTTGGT CTACCCACCA TATGTTTATC AAACTCAGTC
AAGTGCCTTT TGTAGTGTGA AGAAACTCAA AGTAGAGCCA AGCAGTTGTG
TATTCCAGGA AAGAAACTAT CCACGGACCT ATGTGAATGG TAGAAACTTT
GGAAATTCTC ATCCTCCCAC TAAGGGTAGT GCTTTTCAGA CAAAGATACC
ATTTAATAGA CCTCGAGGAC ACAACTTTTC ATTGCAGACA AGTGCTGTTG
TTTTGAAAAA CACTGCAGGT GCTACAAAGG TCATAGCAGC TCAGGCACAG
CAAGCTCACG TGCAGGCACC TCAGATTGGG GCGTGGCGAA ACAGATTGCA
TTTCCTAGAA GGCCCCCAGC GATGTGGATT GAAGCGCAAG AGTGAGGAGT
TGGATAATCA TAGCAGCGCA ATGCAGATTG TCGATGAATT GTCCATACTT
CCTGCAATGT TGCAAACCAA CATGGGAAAT CCAGTGACAG TTGTGACAGC
TACCACAGGA TCAAAACAGA ATTGTACCAC TGGAGAAGGT GACTATCAGT
TAGTACAGCA TGAAGTCTTA TGCTCCATGA AAAATACTTA CGAAGTCCTT
GATTTTCTTG GTCGAGGCAC GTTTGGCCAG GTAGTTAAAT GCTGGAAAAG
AGGGACAAAT GAAATTGTAG CAATCAAAAT TTTGAAGAAT CATCCTTCTT
ATGCCCGTCA AGGTCAAATA GAAGTGAGCA TATTAGCAAG GCTCAGTACT
GAAAATGCTG ATGAATATAA CTTTGTACGA GCTTATGAAT GCTTTCAGCA
CCGTAACCAT ACTTGTTTAG TCTTTGAGAT GCTGGAACAA AACTTGTATG
ACTTTCTGAA ACAAAATAAA TTTAGTCCCC TGCCACTAAA AGTGATTCGG
CCCATTCTTC AACAAGTGGC CACTGCACTG AAAAAATTGA AAAGTCTTGG
TTTAATTCAT GCTGATCTCA AGCCAGAGAA TATTATGTTG GTGGATCCTG
TTCGGCAGCC TTACAGGGTT AAAGTAATAG ACTTTGGGTC GGCCAGTCAT
GTATCAAAGA CTGTTTGTTC AACATATCTA CAATCTCGGT ACTACAGAGC

TCCAGAGATT ATATTGGGGT TGCCATTTTG TGAAGCCATA GACATGTGGT
CATTGGGATG TGTGATTGCA GAATTATTTC TTGGATGGCC GCTCTACCCA
GGAGCCTTGG AGTATGATCA GATTCGATAC ATTTCTCAGA CTCAAGGTTT
GCCAGGAGAA CAGTTGTTAA ATGTGGGTAC TAAATCCACA AGATTTTTTT
GCAAAGAAAC AGATATGTCT CATTCTGGTT GGAGATTAAA GACATTGGAA
GAGCATGAGG CAGAGACAGG AATGAAGTCT AAAGAAGCCA GAAAATACAT
TTTCAACAGT CTGGATGATG TAGCGCATGT GAACACAGTG ATGGATTTGG
AAGGAAGTGA TCTTTTGGCT GAGAAAGCTG ATAGAAGAGA ATTTGTTAGT
CTGTTGAAGA AAATGTTGCT GATTGATGCA GATTTAAGAA TTACTCCAGC
TGAGACCCTG AACCATCCTT TTGTTAATAT GAAACATCTT CTAGATTTCC
CTCATAGCAA CCATGTAAAG TCCTGTTTTC ATATTATGGA TATTTGTAAG
TCCCACCTAA ATTCATGTGA CACAAATAAT CACAACAAAA CTTCACTTTT
AAGACCAGTT GCTTCAAGCA GTACTGCTAC ACTGACTGCA AATTTTACTA
AAATCGGAAC ATTAAGAAGT CAGGCATTGA CCACATCTGC TCATTCAGTT
GTGCACCATG GAATACCTCT GCAGGCAGGA ACTGCTCAGT TTGGTTGTGG
TGATGCTTTT CAGCAGACAT TGATTATCTG TCCCCCAGCT ATTCAAGGTA
TTCCTGCAAC ACATGGTAAA CCCACCAGTT ATTCAATAAG GGTAGATAAT
ACAGTTCCAC TTGTAACTCA GGCCCCAGCT GTGCAGCCAC TACAGATCCG
ACCAGGAGTT CTTTCTCAGA CGTGGTCTGG TAGAACACAG CAGATGCTGG
TGCCTGCCTG GCAACAGGTG ACACCCCTGG CTCCTGCTAC TACTACACTA
ACTTCTGAGA GTGTGGCTGG TTCACACAGG CTTGGAGACT GGGGGAAGAT
GATTTCATGC AGCAATCATT ATAACTCAGT GATGCCGCAG CCTCTTCTGA
CCAATCAGAT AACTTTATCT GCCCCTCAGC CAGTTAGTGT GGGGATTGCA
CATGTTGTCT GGCCTCAGCC TGCCACTACC AAGAAAAATA AACAGTGCCA
GAACAGAGGT ATTTTGGTAA AACTAATGGA ATGGGAGCCA GGAAGAGAGG
AAATAAATGC TTTCAGTTGG AGTAATTCAT TACAGAATAC CAATATCCCA
CATTCAGCAT TTATTTCTCC AAAGATAATT AATGGGAAAG ATGTCGAGGA
AGTAAGTTGT ATAGAAACAC AGGACAATCA GAACTCAGAA GGAGAGGCAA
GAAATTGCTG TGAAACATCT ATCAGACAGG ACTCTGATTC ATCAGTTTCA
GACAAACAGC GGCAAACCAT CATTATTGCC GACTCCCCGA GTCCTGCAGT
GAGTGTCATC ACTATCAGCA GTGACACTGA TGAGGAAGAG ACTTCCCAGA
GACATTCACT CAGAGAATGT AAAGGTAGTC TAGATTGTGA AGCTTGCCAG
AGCACTTTGA ATATTGATCG GATGTGTTCA TTAAGTAGTC CTGATAGTAC
TCTGAGTACC AGCTCCTCAG GGCAGTCCAG CCCATCCCCC TGCAAGAGAC
CGAATAGTAT GTCAGATGAA GAGCAAGAAA GTAGTTGTGA TACGGTGGAT
GGCTCTCCGA CATCTGACTC TTCCGGGCAT GACAGTCCAT TTGCAGAGAG
CACTTTTGTG GAGGACACTC ATGAAAACAC AGAATTGGTA TCCTCTGCTG

ACACAGAAAC CAAGCCAGCT GTCTGTTCTG TTGTGGTGCC ACCAGTGGAA
CTAGAAAATG GCTTAAATGC CGATGAGCAT ATGGCAAACA CAGATTCTAT
ATGCCAGCCA TTAATAAAAG GACGATCTGC CCCTGGAAGA TTAAACCAGC
CTTCTGCAGT GGGTACTCGT CAGCAAAAAT TGACATCAGC ATTCCAGCAG
CAGCATTTGA ACTTCAGTCA GGTTCAGCAC TTTGGATCTG GGCATCAAGA
GTGGAATGGA AACTTTGGGC ACAGAAGACA GCAAGCTTAT ATTCCTACTA
GTGTTACCAG TAATCCATTC ACTCTTTCTC ATGGAAGTCC CAATCACACA
GCAGTGCATG CCCACCTGGC TGGAAATACA CACCTCGGAG GACAGCCTAC
TCTACTTCCA TACCCATCAT CAGCCACCCT CAGTAGTGCT GCACCAGTGG
CCCACCTGTT AGCCTCTCCG TGTACCTCAA GACCTATGTT ACAGCATCCA
ACTTATAATA TCTCCCATCC CAGTGGCATA GTTCACCAAG TCCCAGTGGG
CTTAAATCCC CGTCTGTTAC CATCCCCAAC CATTCATCAG ACTCAGTACA
AACCAATCTT CCCACCACAT TCTTACATTG CAGCATCACC TGCATATACT
GGATTTCCAC TGAGTCCAAC AAAACTCAGC CAGTATCCAT ATATGTGA (SEQ ID No.128)

Protein sequence

MASQVLVYPPYVYQTQSSAFCSVKKLKVEPSSCVFQERNYPRTYVNGRNFGNSHPPTKGS
AFQTKIPFNRPRGHNFSLQTSAVVLKNTAGATKVIAAQAQQAHVQAPQIGAWRNRLHFLEG
PQRCGLKRKSEELDNHSSAMQIVDELSILPAMLQTNMGNPVTVVTATTGSKQNCTTGEGDY
QLVQHEVLCSMKNTYEVLDFLGRGTFGQVVKCWKRGTNEIVAIKILKNHPSYARQGQIEVSI
LARLSTENADEYNFVRAYECFQHRNHTCLVFEMLEQNLYDFLKQNKFSPLPLKVIRPILQQV
ATALKKLKSLGLIHADLKPENIMLVDPVRQPYRVKVIDFGSASHVSKTVCSTYLQSRYYRAPE
IILGLPFCEAIDMWSLGCVIAELFLGWPLYPGALEYDQIRYISQTQGLPGEQLLNVGTKSTRFF
CKETDMSHSGWRLKTLEEHEAETGMKSKEARKYIFNSLDDVAHVNTVMDLEGSDLLAEKA
DRREFVSLLKKMLLIDADLRITPAETLNHPFVNMKHLLDFPHSNHVKSCFHIMDICKSHLNSC
DTNNHNKTSLLRPVASSSTATLTANFTKIGTLRSQALTTSAHSVVHHGIPLQAGTAQFGCGD
AFQQTLIICPPAIQGIPATHGKPTSYSIRVDNTVPLVTQAPAVQPLQIRPGVLSQTWSGRTQQ
MLVPAWQQVTPLAPATTTLTSESVAGSHRLGDWGKMISCSNHYNSVMPQPLLTNQITLSAP
QPVSVGIAHVVWPQPATTKKNKQCQNRGILVKLMEWEPGREEINAFSWSNSLQNTNIPHSA
FISPKIINGKDVEEVSCIETQDNQNSEGEARNCCETSIRQDSDSSVSDKQRQTIIIADSPSPAV
SVITISSDTDEEETSQRHSLRECKGSLDCEACQSTLNIDRMCSLSSPDSTLSTSSSGQSSPS
PCKRPNSMSDEEQESSCDTVDGSPTSDSSGHDSPFAESTFVEDTHENTELVSSADTETKP
AVCSVVVPPVELENGLNADEHMANTDSICQPLIKGRSAPGRLNQPSAVGTRQQKLTSAFQQ
QHLNFSQVQHFGSGHQEWNGNFGHRRQQAYIPTSVTSNPFTLSHGSPNHTAVHAHLAGN
THLGGQPTLLPYPSSATLSSAAPVAHLLASPCTSRPMLQHPTYNISHPSGIVHQ (SEQ ID No.129)

HIPK3 interacts with SnoN and SNIP1, proteins involved in the TGFβ/BMP pathway

By two-hybrid screening in yeast it was shown that PTPN12 interacts with Smad5, a protein involved in the BMP pathway.

SnoN-HIPK3

SID : Nucleic sequence, SEQ ID No.64 and Proteic sequence, SEQ ID No. 102.

Snip1-HIPK3

SID : Nucleic sequence, SEQ ID No.62, 63 and Proteic sequence, SEQ ID No. 100, 101.

Thus, yeast-two-hybrid screens showed that amino-acids 799-1127 from HIPK3 (SEQ ID No.129) interact with amino-acids 1-370 from SnoN (SEQ ID No.26) and that amino-acids 833-930 from HIPK3 interact with amino-acids 1-198 from Snip1 (SEQ ID No.24) (see Fig. 41).

HIPK3 modulates the BMP signaling

Since HIPK3 was found interacting with SnoN and SNIP1, it was investigated whether HIPK3 could be involved in the TGFβ and/or BMP pathways. In order to assay HIPK3's functional involvement in the TGFβ/BMP pathways, HIPK3 cellular knock-down experiments were performed using chemically synthesised siRNA duplexes (see Materials & Methods for siRNA sequences and protocols). While transiently co-transfecting HepG2 cells using the $p(GC)_{12}$-MLP-Luc reporter and HIPK3-targeting siRNA duplex, a specific, dose-dependant and BMP6-dependant increase in the BMP-dependant reporter activity was observed (see Fig. 42 A) demonstrating a function for HIPK3 in the response to the BMP pathway. The positive effect of HIPK3-targeting siRNA duplex was already observed at low concentration of siRNA duplex (4nM) and further enhanced at higher duplex concentrations (40nM). While transiently co-transfecting HepG2 cells using the $p(GTCT)_8$-MLP-Luc reporter and HIPK3-targeting siRNA duplex, no TGFβ-dependant variation in the reporter activity was observed (see Fig. 42 B) demonstrating a restrictive function for HIPK3 on the response the BMP pathway. Modulation of BMP luciferase reporter activities using HIPK3 cellular knock-down shows its implication in the regulation of the BMP pathway.

Examples 18: The following materials and methods were used to obtain the results in examples 8 to 17

18-1.Expression vectors construction

Construction of mammalian baculovirus vector consisted in introduction of mammalian Polymerase II-type transcriptional units such as a promoter active in mammalian cells (for instance CMV, RSV, albumin or inducible promoters). Such plasmids can be used as classical expression vectors to transfect mammalian cells. They can also be used to generate baculoviruses that have the capacity to infect mammalian cells with a high

efficiency where they drive the expression of the gene which is under the transcriptional control of the promoter active in mammalian cells (Kost and Condreay, 2002).

pV3 and pV5 (Figure X) were prepared from pfastbac1 vectors (Invitrogen). First, the *BamHI-EcoRI* fragment of pfastbac1 with the PCR-amplified CMV promoter fragment from pcDNA3.1/Zeo(+) (Invitrogen) was replaced using oligonucleotides oli3054 and oli3055 (PCR conditions as above) to generate pBacCMV.

Oli3054: 5'-cgggatccCGTTGACATTGATTATTGACTAGTT-3' (SEQ ID No.130)

Oli3055: 5'-cggaattcTTGGGTCTCCCTATAGTGAGT-3' (SEQ ID No.131)

Next, the *EcoRI-NotI* fragment of pBacCMV was replaced with the double-stranded oligonucleotide corresponding to the FLAG sequence to generate pBacCMVflag:

5'-AATTCACCATGGATTACAAGGATGACGACGATAAGGC-3' (SEQ ID No.132)

3'-GTGGTACCTAATGTTCCTACTGCTGCTATTCCGCCGG-5' (SEQ ID No.133)

Next, the *XbaI-PstI* fragment of pBacCMVflag was replaced with the double-stranded oligonucleotide to generate pV3:

5'-
CTAGACCCGGGTTGGCCGGACGGGCCCAGCTGCGGCCACTGGGGCCCTTAATTAAGTAA
CTGCA-3' (SEQ ID No.134)

3'-TGGGCCCAACCGGCCTGCCCGGGTCGACGCCGGTGACCCCGGGAATTAATTCATTG-5'
(SEQ ID No.135)

Or the *XbaI-PstI* fragment of pBacCMVflag was replaced with the double-stranded oligonucleotide to generate pV5:

5'-
CTAGACCCGGGTTGGCCGCAGGGGCCCAGCTGCGGCCACTGGGGCCCTTAATTAAGT
AACTGCA-3' (SEQ ID No.136)

3'-
TGGGCCCAACCGGCGTCCCCGGGTCGACGCCGGTGACCCCGGGAATTAATTCATTG-5'
(SEQ ID No.137)

These pV3 and pV5 vectors thus contain a CMV promoter which controls expression of the proteins of interest fused to the FLAG epitope. The MCS, present into pV3 and pV5, contained the SmaI/SfiI/pvuII/SfiI/PacI sites. Differences between pV3 and pV5 were in the SfiI sites:

In pV3, the sfiI sites were oriented whereas in pV5, the SfiI sites were non-oriented Since the preys identified by the two-hybrid assays are cloned between SfiI sites (WO99/42612), the presence of SfiI sites in pV3 and pV5 allowed the direct cloning preys in these mammalian expression vectors.

18-2-1 Gene cloning

Cloning of smad3, RNF11 and LAPTm5 in pV3 was performed by PCR amplification from placenta cDNA library with the following oligonucleotides:

Smad3:

Oli2752: cggactagtCATGTCGTCCATCCTGCCTT (SEQ ID No.138)

Oli2836: gccttaattaaCTAAGACACACTGGAACAGCGG(SEQ ID No.139)

RNF11:

Oli3778: gatcggccggacgggccATGGGGAACTGCCTCAAATCCCCC (SEQ ID No.140)

Oli3779: gatcggccccagtggccTCAATTAGTCTCATAGGATGAAAG (SEQ ID No.141)

LAPTm5:

Oli3776: gatcggccggacgggccATGGACCCCCGCTTGTCCACTGTC (SEQ ID No.142)

Oli3777: gatcggccccagtggccTCACACCTCTGAGTATGGGGGTGG (SEQ ID No.143)

The PCR program was set up as follows:

94°	45"		
94°	45"	+	
55°	45"	¦	X35
72°	7'	+	
72°	7'		
15°			

The amplification was checked by agarose gel electrophoresis. The PCR fragments were purified with Qiaquick column (Qiagen) according to the manufacturer's protocol. The purified PCR fragments were digested with *Sfi*I restriction enzyme (Biolabs) for 1 hour at 50°C. The PCR fragments were purified with Qiaquick column (Qiagen) according to the manufacturer's protocol.

Concerning PP1ca, by the two-hybrid assay a clone was obtained corresponding to nucleotide 1-972 of the PP1ca into the pP6 vector. Concerning FLJ20037, a clone was obtained corresponding to the full-length FLJ20037 cDNA into the pP6 vector. These pP6-PP1ca and pP6-FLJ20037 vectors were digested with *Sfi*I restriction enzyme (Biolabs) for 1 hour at 50°C, extracted, precipitated, and resuspended in water. The PP1ca and FLJ20037 fragments were then purified using Qiaex column (Qiagen) according to the manufacturer's protocol.

18-2-2 Vector preparation

pV3 and pV5 were prepared as previously described (see 4.1).

The pV3 and pV5 vectors were digested with *Sfi*I restriction enzyme (Biolabs) for 1 hour at 50°C, extracted, precipitated, and resuspended in water. Digested plasmid vector backbones were purified on a separation column (Chromaspin TE 400, Clontech) according to the manufacturer's protocol.

18-2-3 Ligation between expression vectors and preys followed by transformation

The digested insert fragments were ligated into an adequately digested (*Sfi*I) vector (pV3 and pV5) according to standard protocol (Sambrook *et al.*) and were transformed into competent bacterial cells. The cells were grown, the DNA extracted and the plasmid was sequenced.

18-2-4 Cell culture

HEK293 cells were propagated in Minimum Essential Medium Eagle (SIGMA) supplemented with 10% fetal bovine serum (FBS, Life technologies, Invitrogen), 100 units ml-1 penicillin, and 100 µg.ml-1 streptomycin (Life Technologies, Invitrogen) at 37°C, 5%CO2 controlled atmosphere.

HepG2 cells were propagated in Dulbecco's modified Eagle's medium (Life Technologies, Invitrogen) supplemented with 10% fetal bovine serum (FBS, Life technologies, Invitrogen), 100 units ml-1 penicillin, and 100 µg.ml-1 streptomycin (Life Technologies, Invitrogen) at 37°C, 5%CO2 controlled atmosphere. Cells were regularly passaged to maintain exponential growth. Twenty four hours before transfections, cells were trypsinized and diluted with fresh medium at 2x105 cells/well in a 24 well plate in order to get approximately 50-80% confluency for transfection.

18-3 Reporters and luciferase assays

3 different reporter vectors : (GTCT)8-MLP-Luc, (CAGA)6-MLP-Luc and (GC)12-MLP-Luc, encoding the firefly luciferase, were generated for luciferase reporter assays. These two first reporters are activated by TGFβ and activins whereas the third one responds to BMPs.

To construct these reporters, the MLP minimal promoter from an adenovirus Major Late gene, containing a TATA box and an initiator element, was first inserted into the BglII and HindIII sites of the pGL3 basic vector (Promega) to generate the MLP-Luc plasmid using the oligonucleotides:

MLP1: 5'-

GATCTGAATTCCATATGCTGCAGGGGCTATAAAAGGGGGTGGGGGCGCGTTCGTCCTC ACTCTCTTCCA-3'(SEQ ID No.144)

and the complementary oligonucleotide

MLP2 : 5'-

AGCTTGGAAGAGAGTGAGGACGAACGCGCCCCCACCCCCTTTTATAGCCCCTGCAGCA TATGGAATTCA-3' (SEQ ID No.145)

To construct $(GTCT)_8$-MLP-Luc, 2 copies of the following annealed oligonucleotides were inserted into the EcoRI site of MLP-Luc. These oligonucleotides contains 4 copies of 'the GTCT box', a TGFβ-responsive sequence (Zawel et al., 1998).

GTCT1 : 5'-AATTCGTCTAGACAAAAGTCTAGACATTTGTCTAGACTAGTGTCTAGACG-3' (SEQ ID No.146)

and the complementary oligonucleotide

GTCT2 : 5'-AATTCGTCTAGACACTAGTCTAGACAAATGTCTAGACTTTTGTCTAGACG-3' (SEQ ID No.147)

To construct $(CAGA)_6$-MLP-Luc, 1 copy of the following annealed oligonucleotides was inserted into the XhoI and NheI sites of MLP-Luc. These oligonucleotides contains 6 copies of 'the CAGA box', a TGF -responsive sequence (Dennler et al., 1998).

CAGA1: 5'-
CTAGAGCCAGACAAAAAGCCAGACATTTAGCCAGACAAAAAGCCAGACATTTAGCCAGA
CAAAAAGCCAGACA-3' (SEQ ID No.148)

and the complementary oligonucleotide

CAGA2: 5'-
TCGATGTGTGGCTTTTTGTCTGGCTAAATGTCTGGCTTTTTGTCTGGCTAAATGTCTGGC
TTTTTGTCTGGCT-3' (SEQ ID No.149)

To construct $(GC)_{12}$-MLP-Luc, 3 copies of the following annealed oligonucleotides were inseted into the XhoI site of MLP-Luc. These oligonucleotides contains 4 copies of 'the GC box', a BMP responsive sequence (Kusanagi et al., 2000).

GC1: 5'- TCGAGCCGCCGCTTTGCCGCCGCTTTGCCGCCGCTTTGCCGCCGC-3' (SEQ ID No.150)

and the complementary oligonucleotide

GC2: 5'- TCGAGCGGCGGCAAAGCGGCGGCAAAGCGGCGGCAAAGCGGCGGC-3' (SEQ ID No.151)

All these constructs were sequence-checked.

These reporters were used to observe the effects of siRNA in transfection experiments (see the siRNA section). These reporters were also used to determine the effect of over-expression of some proteins on TGFβ and/or BMP signaling in co-transfection experiments with pV3 and pV5 vectors encoding proteins of interest (cf expression vectors construction section). To this end, HepG2 and HEK293 cells were transiently transfected using the Fugene 6 (Roche) or the Lipofectamine 2000 (InVitrogen) reagent, respectively, according to the manufacturer recommendations. 400 ng of luciferase reporter and 100 ng of pRL-TK (Promega), encoding the renilla luciferase and used as an internal transfection efficiency control, were transfected per well of a 24 wells-plate. Variable amounts of expression vectors were co-transfectd as indicated in the figures. When increasing amounts of expression vectors were transfected, total DNA was kept constant by the addition of pV3. 24 hours after the transfection, cells were washed and incubated in a medium without serum. 2 hours later, cells were stimulated with 10 ng/mL of human recombinant TGFβ1 (R&D) or 50 ng/mL of human recombinant BMP6 or BMP7 (R&D). 18 to 24 hours after stimulation, Luciferase activities were quantified using the Dual Luciferase reporter assay kit from Promega. Values were normalized with the renilla luciferase activity expressed from pRL-TK.

18-4 Baculovirus infection of mammalian cells

Genetically modified baculoviruses were used to infect mammalian cells (Kost and Condreay, 2002). The pV3 and pV5 mammalian expression vectors, derived from pFastbac1 (see mammalian expression vectors section) and in which the cDNA of genes of interest has been cloned, can be used to produce baculoviruses that can express the protein encoded by this cDNA in mammalian cells. To prepare the baculoviral particles, these vectors were inserted into the baculoviral genome by transposition into *E. coli* competent cells to obtain a recombinant bacmid using the BAC-TO-BAC Baculovirus Expression System (Invitrogen) according to the manufacturer's procedure. Next steps consist in transfecting Sf9 insect cells with this Bacmid DNA and harvesting the viral particles. 'Control' vectors in which the cDNA encoding the β-galactosidase gene or the GFP gene were constructed and used to produce baculoviruses which were expressing the β-galactosidase and GFP proteins. These 'control' baculoviruses allowed to quantify the efficiency of baculovirus infection in mammalian cells by determining the in situ production of β-galactosidase and GFP proteins in cells. Thus, HepG2, HEK293 or HeLa cells were infected with an efficiency higher than 80 % with the mammalian baculoviruses (data not shown).

Transfection of the Sf9 cells with the Bacmid was made with the GeneShuttle reagent (Quantum). The supernatant of the Sf9 cells which contained the recombinant baculoviral particles was harvested 72 hours post transfection. This supernatant can be used to re-infect other Sf9 cells in order to amplify the viral stock in T75 or T150 flasks. In order to check if viral particles have been produced in sufficient amount for the following experiments, viral DNA was quantified by Q-PCR. 600 µL of the Sf9 supernatant, containing the baculoviral particles, supplemented by 1.4 mL of classical cell medium are used per well to infect human cells such as HepG2, HEK 293 or HeLa cells seeded in 24 wells-plate. Then, cellular RNA was extracted 24 to 72 hours post infection to perform Q-PCR experiments (see Q-PCR section). All viruses were conserved at +4°C.

18-5 Quantitative PCR (Q-PCR) experiments

To monitor the biological effects of the proteins of interest in the TGFβ/activin or BMP signaling in cells, quantification of mRNA of genes transcriptionally regulated by TGFβ/activin or BMP by Quantitative-PCR were carried-out using an Applied Biosytems 7000 SDS machine. This quantification follows a transfection of an expression vector of the prey of interest, the transfection of a siRNA or an infection using a genetically-modified baculovirus in mammalian cells such as HepG2, HeLa or HEK 293 cell lines seeded in 24 culture-plate. Cells are then lysed and RNA was extracted using the Rneasy Minikit and the Qia Shredder from Qiagen following the recommendations of the manufacturer. 1 µg of RNA is then used for a reverse transcription reaction to generate the cDNA which will serve as template in the following Q-PCR reaction. The reverse transcription step was realized in 96 wells-plate with

the TaqMan reverse transcription kit (Applied biosystems) following the recommendations of the manufacturer. The cDNA of the gene of interest was then quantified in 96 wells-plate by the SyBR green methodology using the SyBR Green PCR master Mix kit (Applied Bisosystems) in an ABI 7000 machine following the recommendations of the manufacturer. For each reaction, 8 ng of cDNA was used as template and 300 nM of forward and reverse oligonucleotides probing specifically the gene for which the mRNA was quantified are added. Values are normalized with the value obtained for the mRNA of the hGAPDH or hGUS genes which serve as internal experimental controls.

The forward and reverse oligonucleotides probing the gene of interest were designed using the Primer Express software (Applied Biosystems).These oligonucleotides were validated by Q-PCR experiments showing that they allow a quantitative measurement (quantification of cDNA diluted in cascade and PCR efficacy determination).

The human genes used to monitor the effect of TGFβ are the Plasminogen Activator Inhibitor Type 1 gene (hPAI-1) and the JunB gene. The human genes used to monitor the effect of BMPs are the JunB gene and the Alcaline Phosphatase gene (hALP). The genes used as internal quantification controls are the Glyceraldehyde Phosphate Dehydrogenase gene (hGAPDH) and the the β-Glucoronidase gene (hGUS). The sequences of the oligonucleotides probing these mRNA are:

hPAI-1:	forward	TGAAGATCGAGGTGAACGAGAGT (SEQ ID No.152)
	Reverse	GTCCCAGATGAAGGCGTCTTT (SEQ ID No.153)
hJunB:	forward	ACTCATACACAGCTACGGGATACG(SEQ ID No.154)
	Reverse	GGGTCGGCCAGGTTGAC(SEQ ID No.155)
hALP:	forward	CGAGCTGAACAGGAACAACGT(SEQ ID No.156)
	Reverse	CTGCTTGGCTTTTCCTTCATG(SEQ ID No.157)
hGAPDH:	forward	GGAGTCAACGGATTTGGTCGTA(SEQ ID No.158)
	Reverse	GTGGAATCATATTGGAACATGTAAACC(SEQ ID No.159)
hGUS:	forward	CCCGCGGTCGTGATGT(SEQ ID No.160)
	Reverse	TGAGCGATCACCATCTTCAAGT(SEQ ID No.161)

The sequences of the oligonucleotides probing the cDNA of the gene targeted by siRNA and used to validate the effect of the siRNA (see siRNA section) or the over-production level following baculovirus infection were:

ZNF8:	forward	CCAGTCAGGCCATTCCAATT(SEQ ID No.162)
	Reverse	GTGTGCGTTATGGTTAAACGACTTC(SEQ ID No.163)
TβR1:	forward	GTGACTACAACATATTGCTGCAATCAG(SEQ ID No.164)
	Reverse	AGCACACTGGTCCAGCAATG(SEQ ID No.165)
PP1ca :	forward	CTCCACAAGCACGACTTGGA(SEQ ID No.166)
	Reverse	GTTGGGAGCTGAGAAAAGTGTCA(SEQ ID No.167)

KIAA1196: forward GGCCCTCCGAGACATTCC(SEQ ID No.168)
Reverse TAATGGTACTTGAGCCCGTAGATG(SEQ ID No.169)
LMO4: forward CAGAAGGTCTGCTAAAAGGTCAGAGT(SEQ ID No.170)
Reverse GGGATCCACCTGTGATGAACA(SEQ ID No.171)
FLJ20037: forward AACAAAGAGAAGATCACACCACTCA(SEQ ID No.172)
Reverse TAAGACTCCATCGGTCAGAGTCA(SEQ ID No.173)
HYPA : forward TTCCATGCAGCCTGCCTTA(SEQ ID No.174)
Reverse CAGGTGATTTATGTTCAGTCCACAT(SEQ ID No.175)
LAPTm5: forward TGGCCATCTACCATGTGATCA(SEQ ID No.176)
Reverse CGATCCTGAGGTAGCCCATCT(SEQ ID No.177)
HIPK3: forward TTGTTCAACATATCTACAATCTCGGTACT(SEQ ID No.178)
Reverse GAGCGGCCATCCAAGAAATA(SEQ ID No.179)
PTPN12: forward TGTGAGCGCTATTGGCCTTT(SEQ ID No.180)
Reverse TTTTGAAATTCAAGTAAGAGTGTCCTGAT(SEQ ID No.181)

18-6 siRNA

Chemically synthesized siRNA using RNA phosphoramidites were purchased from Genset Oligos /Proligos (Paris, France). siRNA were ordered deprotected, desalted and duplexed.

The siRNA duplexes used in these studies were all 19 ribonucteotides long and contained two thymidines nucleotides at their 3' termini. All siRNA duplexes were designed according to the rules edicted by Tuschl and coll. (Elbashir et al., 2001).

In the following list, all sequences correspond to the sense DNA in the corresponding CDS

- **TβRI**: 5'-GTGTTTCTGCCACCTCTGT-3'(SEQ ID No.182)
- **mTβRI**: 5'-GTGTGTCTGCAACCTCTGT-3'(SEQ ID No.183)
- **PP1ca**: 5'-AACCTTCACTGACTGCTTC-3'(SEQ ID No.184)
- **KIAA1196**: 5'-CGACTGGAAGGATGAGTTC-3'(SEQ ID No.185)
- **HIPK3**: 5'-GCAGTTGTGTATTCCAGGA-3'(SEQ ID No.186)
- **ZNF8**: 5'-GCCTGAAGTCATCTCCCAG-3'(SEQ ID No.187)
- **PTPN12**: 5'-GATATATCCCACAGCCACT-3'(SEQ ID No.188)
- **LMO4**: 5'-GTGGCATGATCCTTTGCAG-3'(SEQ ID No.189)
- **FLJ20037**: 5'-CAAGATCATTGCCACCAGG-3'(SEQ ID No.190)
- **HYPA**: 5'-ATCAATGTGGACTGAACAT-3'(SEQ ID No.191)
- **LAPTm5**: 5'-ATCATGGACTATCTCCTGT-3'(SEQ ID No.192)

As a validation experiment, the efficacy of these siRNA was tested on their targeted mRNA by Q-PCR experiments (see Q-PCR section). Their specificity was assayed on

unrelated mRNA by Q-PCR. All these siRNA inhibit at least from 65 to 95 % the amount of their targeted mRNA and do not show any effect on other unrelated mRNAs.

18-7 SiRNA transfection

Co-transfection of reporter plasmids and siRNAs was carried out with Lipofectamine 2000 (Life Technologies, Invitrogen) as described by the manufacturer for adherent cell lines. Per well 0,4 µg specific luciferase reporter plasmid, 0,1 µg pRL-TK (Promega), 0,5 µg carrier DNA (pBluscript) and 4 to 40 nM siRNA duplex, formulated into liposomes, were applied. The final volume per well was 500µl. Medium were changed 5 hours post-transfection and cells appeared healthy on next day. Cells were serum-starved for 1-2 hours before cytikines treatment. Cells were treated with TGFβ1 (R&D, 5 ng/ml), BMP6 (R&D, 50ng/ml)or BMP7 (R&D, 50ng/ml) for 18 hours before luciferase assay or total RNA extraction.

Luciferase expression was subsequently monitored with the Dual luciferase assay (Promega).

Reporter and carrier plasmids were amplified in DH5β (Stratagene) and purified using the Qiagen Endofree Maxi plasmid Kit.

Transfection of siRNAs for targeting endogenous mRNA was carried out using Oligofectamine (Life Technologies, Invitrogen) and 4 to 40 nM siRNA duplex per well in a 24 well plate.

Specific silencing of targeted genes was confirmed by at least three independent experiments.

18-8 Antibodies

Anti-SARA rabbit polyclonal antibody was purchased from Santa-Cruz (cat # H-300 sc9135) and used at a 1/150 dilution.

Peroxidase-conjugated AffiniPure F(ab')2 fragment donkey anti-Rabbit IgG (H+L) was used as a secondary reagent (1/10000 dilution) and was purchased from Jackson Immunoresearch laboratories, Inc.

18-9 Cell lysis and Immunoblot

Cell were harvested in lysis buffer (2%SDS, 1X PBS), denatured 5 minutes at 95°C and quantified using Bradford reagent (BIORAD) according to the manufacturer's specifications. Cell lysates (20µg/lane) were resolved on a 4-12% NuPAGE gradient gel (Novex, Invitrogen), transfered to 20µm nitrocellulose membrane (Schleicher & Schuell) and blocked in 10% fat-free dried milk in 1X PBS, 0,05% Tween20. Revelation was performed using ECL (Amersham Biosciences) chemoluminescent substrat according to the manufacturer's specifications.

The following results obtained from these Examples, as well as the teachings in the specification are set forth in the Tables below.

While the invention has been described in terms of the various preferred embodiments, the skilled artisan will appreciate that various modifications, substitutions, omissions and changes may be made without departing from the scope thereof. Accordingly, it is intended that the present invention be limited by the scope of the following claims, including equivalents thereof

Table 1 : Bait name and sequence

1-Bait Name	2- Nucleic acid SEQ ID No.	3- Nucleic acid sequence	4 Nucleic position	5. Amino acid SEQ ID No.	6. Amino Acid Sequece
Human Smad1_v2	1	ATGAATGTGACAAGTTTATTTTCCTTTACAAGTCCAGCTGTGAAGAGACTTCTTGGGTGGAAACAGG GCGATGAAGAAGAAAAATGGGCAGAGAAAGCTGTTGATGCTTTGGTGAAAAAACTGAAGAAAAAGAA AGGTGCCATGGAGGAACTGGAAAAGGCCTTGAGCTGCCCAGGGCAACCGAGTAACTGTGTCACCATT CCCCGCTCTCTGGATGGCAGGCTGCAAGTCTCCCACCGGAAGGGACTGCCTCATGTCATTTACTGCC GTGTGTGGCGCTGGCCCGATCTTCAGAGCCACCATGAACTAAAACCACTGGAATGCTGTGAGTTTCC TTTTGGTTCCAAGCAGAAGGAGGTCTGCATCAATCCCTACCACTATAAGAGAGTAGAAAGCCCTGTA CTTCCTCCTGTGCTGGTTCCAAGACACAGCGAATATAATCCTCAGCACAGCCTCTTAGCTCAGTTCC GTAACTTAGGACAAAATGAGCCTCACATGCCACTCAACGCCACTTTTCCAGATTCTTTCCAGCAACC CAACAGCCACCCGTTTCCTCACTCTCCCAATAGCAGTTACCCAAACTCTCCTGGGAGCAGCAGCAGC ACCTACCCTCACTCTCCCACCAGCTCAGACCCAGGAAGCCCTTTCCAGATGCCAGCTGATACGCCCC CACCTGCTTACCTGCCTCCTGAAGACCCCATGACCCAGGATGGCTCTCAGCCGATGGACACAAACAT GATGGCGCCTCCCCTGCCCTCAGAAATCAACAGAGGAGATGTTCAGGCGGTTGCTTATGAGGAAC	[1 802]	14	MNVTSLFSFTSPAVKRLLGWKQG DEEEKWAEKAVDALVKKLKKKK AMEELEKALSCPGQPSNCVTIPR SLDGRLQVSHRKGLPHVIYCRVW RWPDLQSHHELKPLECCEFPFGS KQKEVCINPYHYKRVESPVLPPV LVPRHSEYNPQHSLLAQFRNLGQ NEPHMPLNATFPDSFQQPNSHPF PHSPNSSYPNSPGSSSSTYPHSP TSSDPGSPFQMPADTPPPAYLPP EDPMTQDGSQPMDTNMMAPPLPS EINRGDVQAVAYEE
Human Smad1_v1	2	ATGAATGTGACAAGTTTATTTTCCTTTACAAGTCCAGCTGTGAAGAGACTTCTTGGGTGGAAACAGG GCGATGAAGAAGAAAAATGGGCAGAGAAAGCTGTTGATGCTTTGGTGAAAAAACTGAAGAAAAAGAA AGGTGCCATGGAGGAACTGGAAAAGGCCTTGAGCTGCCCAGGGCAACCGAGTAACTGTGTCACCATT CCCCGCTCTCTGGATGGCAGGCTGCAAGTCTCCCACCGGAAGGGACTGCCTCATGTCATTTACTGCC GTGTGTGGCGCTGGCCCGATCTTCAGAGCCACCATGAACTAAAACCACTGGAATGCTGTGAGTTTCC TTTTGGTTCCAAGCAGAAGGAGGTCTGCATCAATCCCTACCACTATAAGAGAGTAGAAAGCCCTGTA CTTCCTCCTGTGCTGGTTCCAAGACACAGCGAATATAATCCTCAGCACAGCCTCTTAGCTCAGTTCC GTAACTTAGGACAAAATGAGCCTCACATGCCACTCAACGCCACTTTTCCAGATTCTTTCCAGCAACC CAACAGCCACCCGTTTCCTCACTCTCCCAATAGCAGTTACCCAAACTCTCCTGGGAGCAGCAGCAGC ACCTACCCTCACTCTCCCACCAGCTCAGACCCAGGAAGCCCTTTCCAGATGCCAGCTGATACGCCCC CACCTGCTTACCTGCCTCCTGAAGACCCCATGACCCAGGATGGCTCTCAGCCGATGGACACAAACAT GATGGCGCCTCCCCTGCCCTCAGAAATCAACAGAGGAGATGTTCAGGCGGTTGCTTATGAGGAACCA AAACACTGGTGCTCTATTGTCTACTATGAGCTCAACAATCGTGTGGGTGAAGCGTTCCATGCCTCCT CCACAAGTGTGTTGGTGGATGGTTTCACTGATCCTTCCAACAATAAGAACCGTTTCTGCCTTGGGCT GCTCTCCAATGTTAACCGGAATTCCACTATTGAAAACACCAGGCGGCATATTGGAAAAGGAGTTCAT CTTTATTATGTTGGAGGGGAGGTGTATGCCGAATGCCTTAGTGACAGTAGCATCTTTGTGCAAAGTC GGAACTGCAACTACCATCATGGATTTCATCCTACTACTGTTTGCAAGATCCCTAGTGGGTGTAGTCT GAAAATTTTTAACAACCAAGAATTTGCTCAGTTATTGGCACAGTCTGTGAACCATGGATTTGAGACA	[1 1398]	15	MNVTSLFSFTSPAVKRLLGWKQG DEEEKWAEKAVDALVKKLKKKKG AMEELEKALSCPGQPSNCVTIPR SLDGRLQVSHRKGLPHVIYCRVW RWPDLQSHHELKPLECCEFPFGS KQKEVCINPYHYKRVESPVLPPV LVPRHSEYNPQHSLLAQFRNLGQ NEPHMPLNATFPDSFQQPNSHPF PHSPNSSYPNSPGSSSSTYPHSP TSSDPGSPFQMPADTPPPAYLP EDPMTQDGSQPMDTNMMAPPLPS EINRGDVQAVAYEEPKHWCSIVY YELNNRVGEAFHASSTSVLVDGF TDPSNNKNRFCLGLLSNVNRNST IENTRRHIGKGVHLYYVGGEVYA ECLSDSSIFVQSRNCNYHHGFHP TTVCKIPSGCSLKIFNNQEFAQL LAQSVNHGFETVYELTKMCTIRM

		GTCTATGAGCTTACAAAAATGTGTACTATACGTATGAGCTTTGTGAAGGGCTGGGGAGCAGAATACC ACCGCCAGGATGTTACTAGCACCCCCTGCTGGATTGAGATACATCTGCACGGCCCCCTCCAGTGGCT GGATAAAGTTCTTACTCAAATGGGTTCACCTCATAATCCTATTTCATCTGTATCTTAA			SFVKGWGAEYHRQDVTSTPCWIE IHLHGPLQWLDKVLTQMGSPHNP ISSVS
Human Smad1_v3	3	GATGGACACAAACATGATGGCGCCTCCCCTGCCCTCAGAAATCAACAGAGGAGATGTTCAGGCGGTT GCTTATGAGGAACCAAAACACTGGTGCTCTATTGTCTACTATGAGCTCAACAATCGTGTGGGTGAAG CGTTCCATGCCTCCTCCACAAGTGTGTTGGTGGATGGTTTCACTGATCCTTCCAACAATAAGAACCG TTTCTGCCTTGGGCTGCTCTCCAATGTTAACCGGAATTCCACTATTGAAAACACCAGGCGGCATATT GGAAAAGGAGTTCATCTTTATTATGTTGGAGGGGAGGTGTATGCCGAATGCCTTAGTGACAGTAGCA TCTTTGTGCAAAGTCGGAACTGCAACTACCATCATGGATTTCATCCTACTACTGTTTGCAAGATCCC TAGTGGGTGTAGTCTGAAAATTTTTAACAACCAAGAATTTGCTCAGTTATTGGCACAGTCTGTGAAC CATGGATTTGAGACAGTCTATGAGCTTACAAAAATGTGTACTATACGTATGAGCTTTGTGAAGGGCT GGGGAGCAGAATACCACCGCCAGGATGTTACTAGCACCCCCTGCTGGATTGAGATACATCTGCACGG CCCCCTCCAGTGGCTGGATAAAGTTCTTACTCAAATGGGTTCACCTCATAATCCTATTTCATCTGTA TCT	[723 1395]	16	DGHKHDGASPALRNQQRRCSGGC L*GTKTLVLYCLL*AQQSCG*SV PCLLHKCVGGWFH*SFQQ*EPFL PWAALQC*PEFHY*KHQAAYWKR SSSLLCWRGGVCRMP**Q*HLCA KSELQLPSWISSYYCLQDP*WV* SENF*QPRICSVIGTVCEPWI*D SL*AYKNVYYTYELCEGLGSRI PPGCY*HPLLD*DTSARPPPVAG *SSYSNGFTS*SYFICI
Human Smad4_v1	4	ATGGACAATATGTCTATTACGAATACACCAACAAGTAATGATGCCTGTCTGAGCATTGTGCATAGTT TGATGTGCCATAGACAAGGTGGAGAGAGTGAAACATTTGCAAAAAGAGCAATTGAAAGTTTGGTAAA GAAGCTGAAGGAGAAAAAAGATGAATTGGATTCTTTAATAACAGCTATAACTACAAATGGAGCTCAT CCTAGTAAATGTGTTACCATACAGAGAACATTGGATGGGAGGCTTCAGGTGGCTGGTCGGAAAGGAT TTCCTCATGTGATCTATGCCCGTCTCTGGAGGTGGCCTGATCTTCACAAAAATGAACTAAAACATGT TAAATATTGTCAGTATGCGTTTGACTTAAAATGTGATAGTGTCTGTGTGAATCCATATCACTACGAA CGAGTTGTATCACCTGGAATTGATCTCTCAGGATTAACACTGCAGAGTAATGCTCCATCAAGTATGA TGGTGAAGGATGAATATGTGCATGACTTTGAGGGACAGCCATCGTTGTCCACTGAAGGACATTCAAT TCAAACCATCCAGCATCCACCAAGTAATCGTGCATCGACAGAGACATACAGCACCCCAGCTCTGTTA GCCCCATCTGAGTCTAATGCTACCAGCACTGCCAACTTTCCCAACATTCCTGTGGCTTCCACAAGTC AGCCTGCCAGTATACTGGGGGGCAGCCATAGTGAAGGACTGTTGCAGATAGCATCAGGGCCTCAGCC AGGACAGCAGCAGAATGGATTTACTGGTCAGCCAGCTACTTACCATCATAACAGCACTACCACCTGG ACTGGAAGTAGGACTGCACCATACACACCTAATTTGCCTCACCACCAAAACGGCCATCTTCAGCACC ACCCGCCTATGCCGCCCCATCCCGGACATTACTGGCCTGTTCACAATGAGCTTGCATTCCAGCCTCC CATTTCCAATCATCCTGCTCCTGAGTATTGGTGTTCCATTGCTTACTTTGAAATGGATGTTCAGGTA GGAGAGACATTTAAGGTTCCTTCAAGCTGCCCTATTGTTACTGTTGATGGATACGTGGACCCTTCTG GAGGAGATCGCTTTTGTTTGGGTCAACTCTCCAATGTCCACAGGACAGAAGCCATTGAGAGAGCAAG GTTGCACATAGGCAAAGGTGTGCAGTTGGAATGTAAAGGTGAAGGTGATGTTTGGGTCAGGTGCCTT AGTGACCACGCGGTCTTTGTACAGAGTTACTACTTAGACAGAGAAGCTGGGCGTGCACCTGGAGATG CTGTTCATAAGATCTACCCAAGTGCATATATAAAGGTCTTTGATTTGCGTCAGTGTCATCGACAGAT GCAGCAGCAGGCGGCTACTGCACAAGCTGCAGCAGCTGCCCAGGCAGCAGCCGTGGCAGGAAACATC CCTGGCCCAGGATCAGTAGGTGGAATAGCTCCAGCTATCAGTCTGTCAGCTGCTGCTGGAATTGGTG TTGATGACCTTCGTCGCTTATGCATACTCAGGATGAGTTTTGTGAAAGGCTGGGGACCGGATTACCC AAGACAGAGCATCAAAGAAACACCTTGCTGGATTGAAATTCACTTACACCGGGCCCTCCAGCTCCTA GACGAAGTACTTCATACCATGCCGATTGCAGACCCACAACCTTTAGAC	[1 1656]	17	MDNMSITNTPTSNDACLSIVHSL MCHRQGGESETFAKRAIESLVKK LKEKKDELDSLITAITTNGAHPS KCVTIQRTLDGRLQVAGRKGFPH VIYARLWRWPDLHKNELKHVKYC QYAFDLKCDSVCVNPYHYERVVS PGIDLSGLTLQSNAPSSMMVKDE YVHDFEGQPSLSTEGHSIQTIQH PPSNRASTETYSTPALLAPSESN ATSTANFPNIPVASTSQPASILG GSHSEGLLQIASGPQPGQQQNGF TGQPATYHHNSTTTWTGSRTAPY TPNLPHHQNGHLQHHPPMPPHPG HYWPVHNELAFQPPISNHPAPEY WCSIAYFEMDVQVGETFKVPSSC PIVTVDGYVDPSGGDRFCLGQLS NVHRTEAIERARLHIGKGVQLE KGEGDVWVRCLSDHAVFVQSYYL DREAGRAPGDAVHKIYPSAYIKV FDLRQCHRQMQQQAATAQAAAAA QAAAVAGNIPGPGSVGGIAPAIS LSAAAGIGVDDLRRLCILRMSFV KGWGPDYPRQSIKETPCWIEIHL HRALQLLDEVLHTMPIADPQPLD

WO 03/045990 PCT/EP02/13866

Human Smad4_v3	5	GAATGGATTTACTGGTCAGCCAGCTACTTACCATCATAACAGCACTACCACCTGGACTGGAAGTAGG ACTGCACCATACACACCTAATTTGCCTCACCACCAAAACGGCCATCTTCAGCACCACCCGCCTATGC CGCCCCATCCCGGACATTACTGGCCTGTTCACAATGAGCTTGCATTCCAGCCTCCCATTTCCAATCA TCCTGCTCCTGAGTATTGGTGTTCCATTGCTTACTTTGAAATGGATGTTCAGGTAGGAGAGACATTT AAGGTTCCTTCAAGCTGCCCTATTGTTACTGTTGATGGATACGTGGACCCTTCTGGAGGAGATCGCT TTTGTTTGGGTCAACTCTCCAATGTCCACAGGACAGAAGCCATTGAGAGAGCAAGGTTGCACATAGG CAAAGGTGTGCAGTTGGAATGTAAAGGTGAAGGTGATGTTTGGGTCAGGTGCCTTAGTGACCACGCG GTCTTTGTACAGAGTTACTACTTAGACAGAGAAGCTGGGCGTGCACCTGGAGATGCTGTTCATAAGA TCTACCCAAGTGCATATATAAAGGTCTTTGATTTGCGTCAGTGTCATCGACAGATGCAGCAGCAGGC GGCTACTGCACAAGCTGCAGCAGCTGCCCAGGCAGCAGCCGTGGCAGGAAACATCCCTGGCCCAGGA TCAGTAGGTGGAATAGCTCCAGCTATCAGTCTGTCAGCTGCTGCTGGAATTGGTGTTGATGACCTTC GTCGCTTATGCATACTCAGGATGAGTTTTGTGAAAGGCTGGGGACCGGATTACCCAAGACAGAGCAT CAAAGAAACACCTTGCTGGATTGAAATTCACTTACACCGGGCCCTCCAGCTCCTAGACGAAGTACTT CATACCATGCCGATTGCAGACCCACAACCTTTAGAC	[750 1656]	18	EWIYWSASYLPS*QHYHLDWK*D CTIHT*FASPPKRPSSAPPAYAA PSRTLLACSQ*ACIPASHFQSSC S*VLVFHCLL*NGCSGRRDI*GS FKLPYCYC*WIRGPFWRRSLLFG STLQCPQDRSH*ESKVAHRQRCA VGM*R*R*CLGQVP**PRGLCTE LLLRQRSWACTWRCCS*DLPKCI YKGL*FASVSSTDAAAGGYCTSC SSCPGSSRGRKHPWPRISRWNSS SYQSVSCCWNWC**PSSLMHTQ EFCERLGTGLPKTEHQRNTLLD* NSLTPGPPAPRRSTSYHADCRPT TFR
Human Smad5_v2	6.	ATGACGTCAATGGCCAGCTTGTTTTCTTTTACTAGTCCAGCAGTAAAGCGATTGTTGGGCTGGAAAC AAGGTGATGAGGAGGAGAAATGGGCAGAAAAGGCAGTTGATGCTTTGGTGAAGAAACTAAAAAAGAA AAAGGGTGCCATGGAGGAACTGGAGAAAGCCTTGAGCAGTCCAGGACAGCCGAGTAAATGTGTCACT ATTCCCAGATCTTTAGATGGACGCCTGCAGGTTTCTCACAGAAAAGGCTTACCCCATGTTATATATT GTCGTGTTTGGCGCTGGCCGGATTTGCAGAGTCATCATGAGCTAAAGCCGTTGGATATTTGTGAATT TCCTTTTGGATCTAAGCAAAAAGAAGTTTGTATCAACCCATACCACTATAAGAGAGTGGAGAGTCCA GTCTTACCTCCAGTATTAGTGCCTCGTCATAATGAATTCAATCCACAACACAGCCTTCTGGTTCAGT TTAGGAACCTGAGCCACAATGAACCACACATGCCACAAAATGCCACGTTTCCACATTCTTTCCACCA GCCCAACAACACTCCTTTTCCCTTATCTCCAAACAGCCCTTATCCCCCTTCTCCTGCTAGCAGCACA TATCCCAACTCCCCAGCAAGTTCTGGACCAGGAAGTCCATTTCAGCTCCCAGCTGATACGCCTCCTC CTGCCTATATGCCACCTGATGATCAGATGGGTCAAGATCCTTCCCAGCCTATGGATACAAGCAATAA TATGATTCCTCAGATTATGCCCAGTATATCCAGCAGGGATGTTCAGCCTGTTGCCTATGAAGAGCC	[1 803]	19	MTSMASLFSFTSPAVKRLLGWKQ GDEEEKWAEKAVDALVKKLKKKK GAMEELEKALSSPGQPSKCVTIP RSLDGRLQVSHRKGLPHVIYCRV WRWPDLQSHHELKPLDICEFPFG SKQKEVCINPYHYKRVESPVLPP VLVPRHNEFNPQHSLLVQFRNLS HNEPHMPQNATFPHSFHQPNNTP FPLSPNSPYPPSPASSTYPNSPA SSGPGSPFQLPADTPPPAYMPPD DQMGQDPSQPMDTSNNMIPQIMP SISSRDVQPVAYEE
Smad9a isoform_v1	7	ATGCACTCCACCACCCCCATCAGCTCCCTCTTCTCCTTCACCAGCCCCGCAGTGAAGAGACTGCTAG GCTGGAAGCAAGGAGATGAAGAGGAAAAGTGGGCAGAGAAGGCAGTGGACTCTCTAGTGAAGAAGTT AAAGAAGAAGAAGGGAGCCATGGACGAGCTGGAGAGGGCTCTCAGCTGCCCGGGGCAGCCCAGCAAA TGCGTCACGATTCCCCGCTCCCTGGACGGGCGGCTGCAGGTGTCCCACCGCAAGGGCCTGCCCCATG TGATTTACTGTCGCGTGTGGCGCTGGCCGGATCTGCAGTCCCACCACGAGCTGAAGCCGCTGGAGTG CTGTGAGTTCCCATTTGGCTCCAAGCAGAAAGAAGTGTGCATTAACCCTTACCACTACCGCCGGGTG GAGACTCCAGTACTGCCTCCTGTGCTCGTGCCAAGACACAGTGAATATAACCCCCAGCTCAGCCTCC TGGCCAAGTTCCGCAGCGCCTCCCTGCACAGTGAGCCACTCATGCCACACAACGCCACCTATCCTGA CTCTTTCCAGCAGCCTCCGTGCTCTGCACTCCCTCCCTCACCCAGCCACGCGTTCTCCCAGTCCCCG TGCACGGCCAGCTACCCTCACTCCCCAGGAAGTCCTTCTGAGCCAGAGAGTCCCTATCAACACTCAG TTGACACACCACCCCTGCCTTATCATGCCACAGAAGCCTCTGAGACCCAGAGTGGCCAACCTGTAGA TGCCACAGCTGATAGACATGTAGTGCTATCGATACCAAATGGAGACTTTCGACCAGTTTGTTACGAG GAGCCCCAGCACTGGTGCTCGGTCGCCTACTATGAACTGAACAACCGAGTTGGGGAGACATTCCAGG	[1 1404]	20	MHSTTPISSLFSFTSPAVKRLLG WKQGDEEKWAEKAVDSLVKKLK KKKGAMDELERALSCPGQPSKCV TIPRSLDGRLQVSHRKGLPHVIY CRVWRWPDLQSHHELKPLECCE PFGSKQKEVCINPYHYRRVETPV LPPVLVPRHSEYNPQLSLLAKFR SASLHSEPLMPHNATYPDSFQQP PCSALPPSPSHAFSQSPCTASYP HSPGSPSEPESPYQHSVDTPPLP YHATEASETQSGQPVDATADRHV VLSIPNGDFRPVCYEEPQHWCSV AYYELNNRVGETFQASSRSVLID

		CTTCCTCCCGAAGTGTGCTCATAGATGGGTTCACCGACCCTTCAAATAACAGGAACAGATTCTGTCT TGGACTTCTTTCTAATGTAAACAGAAACTCAACGATAGAAAATACCAGGAGACATATAGGAAAGGGT GTGCACTTGTACTACGTCGGGGGAGAGGTGTATGCCGAGTGCGTGAGTGACAGCAGCATCTTTGTGC AGAGCCGGAACTGCAACTATCAACACGGCTTCCACCCAGCTACCGTCTGCAAGATCCCCAGCGGCTG CAGCCTCAAGGTCTTCAACAACCAGCTCTTCGCTCAGCTCCTGGCCCAGTCAGTTCACCACGGCTTT GAAGTCGTGTATGAACTGACCAAGATGTGTACTATCCGGATGAGTTTTGTTAAGGGTTGGGGTGCTG AGTATCATCGCCAGGATGTCACCAGCACCCCCTGCTGGATTGAGATTCATCTTCATGGGCCACTGCA GTGGCTGGACAAAGTTCTGACTCAGATGGGCTCTCCACATAACCCCATTTCTTCAGTGTCTTAA			GFTDPSNNRNRFCLGLLSNVNRN STIENTRRHIGKGVHLYYVGGEV YAECVSDSSIFVQSRNCNYQHGF HPATVCKIPSGCSLKVFNNQLFA QLLAQSVHHGFEVVYELTKMCTI RMSFVKGWGAEYHRQDVTSTPCW IEIHLHGPLQWLDKVLTQMGSPH NPISSVS
Human Smad9_v3	8	CTCCCCAGGAAGTCCTTCTGAGCCAGAGAGTCCCTATCAACACTCAGACTTTCGACCAGTTTGTTAC GAGGAGCCCCAGCACTGGTGCTCGGTCGCCTACTATGAACTGAACAACCGAGTTGGGGAGACATTCC AGGCTTCCTCCCGAAGTGTGCTCATAGATGGGTTCACCGACCCTTCAAATAACAGGAACAGATTCTG TCTTGGACTTCTTTCTAATGTAAACAGAAACTCAACGATAGAAAATACCAGGAGACATATAGGAAAG GGTGTGCACTTGTACTACGTCGGGGGAGAGGTGTATGCCGAGTGCGTGAGTGACAGCAGCATCTTTG TGCAGAGCCGGAACTGCAACTATCAACACGGCTTCCACCCAGCTACCGTCTGCAAGATCCCCAGCGG CTGCAGCCTCAAGGTCTTCAACAACCAGCTCTTCGCTCAGCTCCTGGCCCAGTCAGTTCACCACGGC TTTGAAGTCGTGTATGAACTGACCAAGATGTGTACTATCCGGATGAGTTTTGTTAAGGGTTGGGGTG CTGAGTATCATCGCCAGGATGTCACCAGCACCCCCTGCTGGATTGAGATTCATCTTCATGGGCCACT GCAGTGGCTGGACAAAGTTCTGACTCAGATGGGCTCTCCACATAACCCCATTTCTTCAGTGTCTTAA	[624 1293]	21	LPRKSF*ARESLSTLRLSTSLLR GAPALVLGRLL*TEQPSWGDIPG FLPKCAHRWVHRPFK*QEQILS TSF*CKQKLNDRKYQETYRKGCA LVLRRGRGVCRVRE*QQHLCAEP ELQLSTRLPPSYRLQDPQRLQPQ GLQQPALRSAPGPVSSPRL*SRV *TDQDVYYPDEFC*GLGC*VSSP GCHQHPLLD*DSSSWATAVAGQS SDSDGLST*PHFFSVL
Human SMURF2_v2	9	ATGTCTAACCCCGGAGGCCGGAGGAACGGGCCCGTCAAGCTGCGCCTGACAGTACTCTGTGCAAAAA ACCTGGTGAAAAAGGATTTTTTCCGACTTCCTGATCCATTTGCTAAGGTGGTGGTTGATGGATCTGG GCAATGCCATTCTACAGATACTGTGAAGAATACGCTTGATCCAAAGTGGAATCAGCATTATGACCTG TATATTGGAAAGTCTGATTCAGTTACGATCAGTGTATGGAATCACAAGAAGATCCATAAGAAACAAG GTGCTGGATTTCTCGGTTGTGTTCGTCTTCTTTCCAATGCCATCAACCGCCTCAAAGACACTGGTTA TCAGAGGTTGGATTTATGCAAACTCGGGCCAAATGACAATGATACAGTTAGAGGACAGATAGTAGTA AGTCTTCAGTCCAGAGACCGAATAGGCACAGGAGGACAAGTTGTGGACTGCAGTCGTTTATTTGATA ACGATTTACCAGACGGCTGGGAAGAAAGGAGAACCGCCTCTGGAAGAATCCAGTATCTAAACCATAT AACAAGAACTACGCAATGGGAGCGCCCAACACGACCGGCATCCGAATATTCTAGCCCTGGCAGACCT CTTAGCTGCTTTGTTGATGAGAACACTCCAATTAGTGGAACAAATGGTGCAACATGTGGACAGTCTT CAGATCCCAGGCTGGCAGAGAGGAGAGTCAGGTCACAACGACATAGAAATTACATGAGCAGAACACA TTTACATACTCCTCCAGACCTACCAGAAGGCTATGAACAGAGGACAACGCAACAAGGCCAGGTGTAT TTCTTACATACACAGACTGGTGTGAGCACATGGCATGATCCAAGAGTGCCCAGGGATCTTAGCAACA TCAATTGTGAAGAGCTTGGTCCGTTGCCTCCTGGATGGGAGATCCGTAATACGGCAACAGGCAGAGT TTATTTCGTTGACCATAACAACAGAACAACACAATTTACAGATCCTCGGCTGTCTGCTAACTTGCAT	[1 1005]	22	MSNPGGRRNGPVKLRLTVLCAKN LVKKDFFRLPDPFAKVVVDGSGQ CHSTDTVKNTLDPKWNQHYDLYI GKSDSVTISVWNHKKIHKKQGAG FLGCVRLLSNAINRLKDTGYQRL DLCKLGPNDNDTVRGQIVVSLQS RDRIGTGGQVVDCSRLFDNDLPD GWEERRTASGRIQYLNHITRTTQ WERPTRPASEYSSPGRPLSCFVD ENTPISGTNGATCGQSSDPRLAE RRVRSQRHRNYMSRTHLHTPPDL PEGYEQRTTQQGQVYFLHTQTGV STWHDPRVPRDLSNINCEELGP PPGWEIRNTATGRVYFVDHNNRT TQFTDPRLSANLH
Human SARA_v5	10	AGTGCCTCAAGCCAGAGCCCTAACCCTAACAATCCTGCTGAATACTGTTCTACTATCCCTCCCTTGC AGCAAGCTCAGGCCTCAGGAGCTCTGAGCTCTCCACCTCCCACTGTGATGGTACCTGTGGGAGTTTT AAAGCACCCTGGAGCAGAAGTGGCTCAGCCCAGAGAGCAGAGGCGAGTTTGGTTTGCTGATGGGATC TTGCCCAATGGAGAAGTTGCTGATGCAGCCAAATTAACAATGAATGGAACTTCCTCTGCAGGAACCC TGGCTGTGTCACACGACCCAGTCAAGCCAGTAACTACCAGTCCTCTACCAGCAGAGACGGATATTTG TCTATTCTCTGGGAGTATAACTCAGGTTGGAAGTCCTGTTGGAAGTGCAATGAATCTTATTCCTGAA	[1996 3972]	23	SASSQSPNPNNPAEYCSTIPPLQ QAQASGALSSPPPTVMVPVGVLK HPGAEVAQPREQRRVWFADGILP NGEVADAAKLTMNGTSSAGTLAV SHDPVKPVTTSPLPAETDICLFS GSITQVGSPVGSAMNLIPEDGLP

		GATGGCCTTCCTCCCATTCTCATCTCCACTGGTGTAAAAGGAGACTATGCTGTGGAAGAGAAACCAT CACAGATTTCAGTAATGCAGCAGTTGGAGGATGGTGGCCCTGACCCACTTGTATTTGTTTTAAATGC AAATTTGTTGTCAATGGTTAAAATTGTAAATTATGTGAACAGGAAGTGCTGGTGTTTCACAACCAAG GGAATGCATGCAGTGGGTCAGTCTGAGATAGTCATTCTTCTACAGTGTTTACCGGATGAAAAGTGTT TGCCAAAGGATATCTTTAATCACTTTGTGCAGCTTTATCGGGATGCTCTGGCAGGGAATGTGGTGAG CAACTTGGGACATTCCTTCTTCAGTCAAAGTTTCCTTGGCAGTAAAGAACATGGTGGATTCTTATAT GTGACATCTACCTACCAGTCACTGCAAGACCTAGTACTCCCAACCCCACCTTACTTGTTTGGGATTC TTATCCAGAAATGGGAAACTCCTTGGGCTAAAGTATTTCCTATCCGTCTGATGTTGAGACTTGGAGC TGAATATCGACTTTATCCATGCCCACTATTCAGTGTCAGATTTCGGAAGCCATTGTTTGGAGAGACG GGGCATACCATCATGAATCTTCTTGCAGACTTCAGAAATTACCAGTATACCTTGCCAGTAGTTCAAG GTTTGGTGGTTGATATGGAAGTTCGGAAAACTAGCATCAAAATTCCCAGCAACAGATACAATGAGAT GATGAAAGCCATGAACAAGTCCAATGAGCATGTCCTGGCAGGAGGTGCCTGCTTCAATGAAAAGGCA GACTCTCATCTTGTGTGTGTACAGAATGATGATGGAAACTATCAGACCCAGGCTATCAGTATTCACA ATCAGCCCAGAAAAGTGACTGGTGCCAGTTTCTTTGTGTTCAGTGGCGCTCTGAAATCCTCTTCTGG ATACCTTGCCAAGTCCAGTATTGTGGAAGATGGTGTTATGGTCCAGATTACTGCAGAGAACATGGAT TCCTTGAGGCAGGCACTGCGAGAGATGAAGGACTTCACCATCACCTGTGGGAAGGCGGACGCGGAGG AACCCCAGGAGCACATCCACATCCAGTGGGTGGATGATGACAAGAACGTTAGCAAGGGTGTCGTAAG TCCTATAGATGGGAAGTCCATGGAGACTATAACAAATGTGAAGATATTCCATGGATCAGAATATAAA GCAAATGGAAAAGTAATCAGATGGACAGAGGTGTTTTTCCTAGAAAACGATGACCAGCACAATTGCC TCAGTGATCCTGCAGATCACAGTAGATTGACTGAGCATGTTGCCAAAGCTTTTTGCCTTGCTCTCTG TCCTCACCTGAAACTTCTGAAGGAAGATGGAATGACCAAACTGGGACTACGTGTGACACTTGACTCA GATCAGGTTGGCTATCAAGCAGGGAGCAATGGCCAGCCCCTTCCCTCGCAGTACATGAATGATCTGG ATAGCGCCTTGGTGCCGGTGATCCATGGAGGGGCCTGCCAGCTTAGTGAGGGCCCCGTTGTCATGGA ACTCATCTTTTATATTCTGGAAAACATCGTATAA			PILISTGVKGDYAVEEKPSQISV MQQLEDGGPDPLVFVLNANLLSM VKIVNYVNRKCWCFTTKGMHAVG QSEIVILLQCLPDEKCLPKDIFN HFVQLYRDALAGNVVSNLGHSFF SQSFLGSKEHGGFLYVTSTYQSL QDLVLPTPPYLFGILIQKWETPW AKVFPIRLMLRLGAEYRLYPCPL FSVRFRKPLFGETGHTIMNLLAD FRNYQYTLPVVQGLVVDMEVRKT SIKIPSNRYNEMMKAMNKSNE LAGGACFNEKADSHLVCVQNDDG NYQTQAISIHNQPRKVTGASFFV FSGALKSSSGYLAKSSIVEDGVM VQITAENMDSLRQALREMKDFTI TCGKADAEEPQEHIHIQWVDDDK NVSKGVVSPIDGKSMETITNVKI FHGSEYKANGKVIRWTEVFFLEN DDQHNCLSDPADHSRLTEHVAKA FCLALCPHLKLLKEDGMTKLGLR VTLDSDQVGYQAGSNGQPLPSQY MNDLDSALVPVIHGGACQLSEGP VVMELIFYILENIV
hSNIP1_F1	11	ATGAAGGCGGTGAAGAGCGAACGGGAGCGAGGGAGCCGGCGAAGACACCGGGACGGGGACGTGGTGC TGCCGGCGGGGGTGGTAGTGAAGCAGGAGCGTCTCAGCCCAGAAGTCGCACCTCCCGCCCACCGCCG TCCGGACCACTCCGGTGGTAGCCCGTCTCCGCCGACCAGCGAGCCGGCCCGCTCGGGCCACCGCGGG AACCGAGCCCGAGGAGTTAGCCGGTCCCCACCCAAAAAGAAAAACAAGGCCTCAGGGAGAAGAAGCA AGTCTCCTCGCAGTAAGAGAAACCGAAGTCCTCACCACTCAACAGTCAAAGTGAAGCAGGAGCGTGA GGATCATCCCCGGAGAGGACGGGAGGATCGGCAGCACAGGGAACCATCAGAACAGGAACACAGGAGA GCTAGGAACAGTGACCGGGACAGACACCGGGGCCATTCCCACCAAAGGAGAACGTCTAACGAGAGGC CTGGGAGTGGGCAGGGTCAGGGACGGGATCGAGACACTCAGAACCTGCAGGCTCAGGAAGAAGAGCG GGAGTTTTATAATGCCAGGCGACGGGAGCATCGCCAGAGGAATGACGTTGGTGGTGGCGGCAGTGAG TCTCAGGAGTTGGTTCCTCGGCCTGGTGGCAACAACAAAGAAAAAGAGGTGCCCGCTAAAGAAAAAC CAAGCTTTGAACTTTCTGGGGCACTTCTTGAGGACACCAACACTTTCCGGGGTGTAGTCATTAAATA TAGTGAGCCCCCAGAAGCACGTATCCCCAAAAAACGGTGGCGTCTCTACCCATTTAAAAATGATGAG GTGCTTCCAGTCATGTACATACATCGACAGAGTGCGTACCTACTGGGTCGACACCGCCGCATTGCAG ACATTCCAATTGATCACCCGTCTTGTTCAAAGCAGCATGCGGTCTTTCAATATCGGCTTGTGGAATA TACCCGTGCTGATGGCACAGTTGGCCGAAGAGTGAAGCCCTACATCATTGACCTTGGCTCAGGCAAT	[1 - 1188]	24	MKAVKSERERGSRRRHRDGDVVL PAGVVVKQERLSPEVAPPAHRRP DHSGGSPSPPTSEPARSGHRGNR ARGVSRSPPKKKNKASGRRSKSP RSKRNRSPHHSTVKVKQEREDHP RRGREDRQHREPSEQEHRRARN DRDRHRGHSHQRRTSNERPGS GQGRDRDTQNLQAQEEEREFYNA RRREHRQRNDVGGGGSESQELVP RPGGNNKEKEVPAKEKPSFELSG ALLEDTNTFRGVVIKYSEPPEAF IPKKRWRLYPFKNDEVLPVMYIP RQSAYLLGRHRRIADIPIDHPSC SKQHAVFQYRLVEYTRADGTVGF RVKPYIIDLGSGNGTFLNNKRIE

		GGAACCTTCTTAAACAACAAACGTATTGAGCCACAGAGATACTATGAACTAAAAGAAAAGGATGTAC TCAAATTTGGATTCAGTATCAGAGAATACGTCTTGCTCCATGAGTCGTCGGACACTTCTGAAATAGA CAGGAAAGATGACGAGGATGAGGAGGAGGAGGAAGAAGTGTCTGACAGC			PQRYYELKEKDVLKFGFSIREYV LLHESSDTSEIDRKDDEDEEEEE EVSDS
hSNIP1_F2	12	ATGAAGGCGGTGAAGAGCGAACGGGAGCGAGGGAGCCGGCGAAGACACCGGGACGGGGACGTGGTGC TGCCGGCGGGGGTGGTAGTGAAGCAGGAGCGTCTCAGCCCAGAAGTCGCACCTCCCGCCCACCGCCG TCCGGACCACTCCGGTGGTAGCCCGTCTCCGCCGACCAGCGAGCCGGCCCGCTCGGGCCACCGCGGG AACCGAGCCCGAGGAGTTAGCCGGTCCCCACCCAAAAAGAAAAACAAGGCCTCAGGGAGAAGAAGCA AGTCTCCTCGCAGTAAGAGAAACCGAAGTCCTCACCACTCAACAGTCAAAGTGAAGCAGGAGCGTGA GGATCATCCCCGGAGAGGACGGGAGGATCGGCAGCACAGGGAACCATCAGAACAGGAACACAGGAGA GCTAGGAACAGTGACCGGGACAGACACCGGGGCCATTCCCACCAAAGGAGAACGTCTAACGAGAGGC CTGGGAGTGGGCAGGGTCAGGGACGGGATCGAGACACTCAGAACCTGCAGGCTCAGGAAGAAGAGCG GGAGTTTTATAATGCCAGGCGACGGGAGCATCGCCAGAGGAATGACGTTGGTGGTGGC	[1-594]	25	MKAVKSERERGSRRRHRDGDVVL PAGVVVKQERLSPEVAPPAHRRP DHSGGSPSPPTSEPARSGHRGNR ARGVSRSPPKKKNKASGRRSKSP RSKRNRSPHHSTVKVKQEREDHP RRGREDRQHREPSEQEHRRARNS DRDRHRGHSHQRRTSNERPGSGQ GQGRDRDTQNLQAQEEEREFYN RRREHRQRNDVGGG
Human hSnoN_v2	13	ATGGAAAACCTCCAGACAAATTTCTCCTTGGTTCAGGGCTCAACTAAAAAACTGAATGGGATGGGAG ATGATGGCAGCCCCCCAGCGAAAAAAATGATAACGGACATTCATGTAAATGGAAAAACGATAAACAA GGTGCCAACAGTTAAGAAGGAACACTTGGATGACTATGGAGAAGCACCAGTGGAAACTGATGGAGAG CATGTTAAGCGAACCTGTACTTCTGTTCCTGAAACTTTGCATTTAAATCCCAGTTTGAAACACACAT TGGCACAATTCCATTTAAGTAGTCAGAGCTCGCTGGGTGGACCAGCAGCATTTTCTGCTCGGCATTC CCAAGAAAGCATGTCGCCTACTGTATTTCTGCCTCTTCCATCACCTCAGGTTCTTCCTGGCCCATTG CTCATCCCTTCAGATAGCTCCACAGAACTCACTCAGACTGTGTTGGAAGGGGAATCTATTTCTTGTT TTCAAGTTGGAGGAGAAAAGAGACTCTGTTTGCCCCAAGTCTTAAATTCTGTTCTCCGAGAATTTAC ACTCCAGCAAATAAATACAGTGTGTGATGAACTGTACATATATTGTTCAAGGTGTACTTCAGACCAG CTTCATATCTTAAAGGTACTGGGCATACTTCCATTCAATGCCCCATCCTGTGGGCTGATTACATTAA CTGATGCACAAAGATTATGTAATGCTTTATTGCGGCCACGAACTTTTCCTCAAAATGGTAGCGTACT TCCTGCTAAAAGCTCATTGGCCCAGTTAAAGGAAACTGGCAGTGCCTTTGAAGTGGAGCATGAATGC CTAGGCAAATGTCAGGGTTTATTTGCACCCCAGTTTTATGTTCAGCCTGATGCTCCGTGTATTCAAT GTCTGGAGTGTTGTGGAATGTTTGCACCCCAGACGTTTGTGATGCATTCTCACAGATCACCTGACAA AAGAACTTGCCACTGGGGCTTTGAATCAGCTAAATGGCATTGCTATCTTCATGTGAACCAAAAATAC TTAGGAACACCTGAAGAAAAGAAACTGAAGATAATTTTAGAAGAAATGAAGGAGAAGTTTAGCATGA GAAGTGGAAAGAGAAATCAATCCAAGACAGATGCACCA	[1 1110]	26	MENLQTNFSLVQGSTKKLNGMGD DGSPPAKKMITDIHVNGKTINKV PTVKKEHLDDYGEAPVETDGEHV KRTCTSVPETLHLNPSLKHTLAQ FHLSSQSSLGGPAAFSARHSQES MSPTVFLPLPSPQVLPGPLLIPS DSSTELTQTVLEGESISCFQVGG EKRLCLPQVLNSVLREFTLQQIN TVCDELYIYCSRCTSDQLHILKV LGILPFNAPSCGLITLTDAQRLC NALLRPRTFPQNGSVLPAKSSLA QLKETGSAFEVEHECLGKCQGLF APQFYVQPDAPCIQCLECCGMFA PQTFVMHSHRSPDKRTCHWGFES AKWHCYLHVNQKYLGTPEEKKLK IILEEMKEKFSMRSGKRNQSKTD AP

WO 03/045990 PCT/EP02/13866

Table 2 : Bait-prey interactions

1: Bait name	2: Bait nucleic acid SEQ ID No.	3: Bait construction	4: Prey name	5: Prey construction
Human Smad1_v2	1	pB27	prey000356 - Human ZNF8	Placenta Random Primed 2
Human Smad1_v2	1	pB27	hgx554 (ZNF8 HF.18; prey69489) hZNF8	Human Breast Epithelial cells RP1
Human Smad1_v2	1	pB27	hgx554 (ZNF8 HF.18; prey69489) hZNF8	Human CEMC7 Random Primed
Human Smad1_v2	1	pB27	prey156850 (ZNF8 HF.18) hZNF8	Human Thymocytes Random Primed 1
Human Smad1_v1	2	pB27	prey69489 (ZNF8) hZNF8	Human Placenta Random Primed 2
Human Smad1_v1	2	pB27	hgx554 (ZNF8 HF.18; prey69489) hZNF8	Human Breast Epithelial cells RP1
Human Smad1_v1	2	pB27	hgx554 (ZNF8 HF.18; prey69489) hZNF8	Human Thymocytes Random Primed 1
Human Smad1_v1	2	pB27	hgx554 (ZNF8 HF.18; prey69489) hZNF8	Human CEMC7 Random Primed
Human Smad1_v1	2	pB27	prey17409 (KIAA1196) hKIAA1196	Human Placenta Random Primed 2
Human Smad1_v1	2	pB27	hgx559 (KIAA1196; prey106053) hKIAA1196	Human Breast Epithelial cells RP1
Human Smad1_v3	3	pB27	prey106053 (KIAA1196) hKIAA1196	Human Placenta Random Primed 2
Human Smad4_v1	4	pB27	prey69401 (prey159865) hZNF83	Human Thymocytes Random Primed 1
Human Smad4_v3	5	pB27	prey024113 - Human HYPA	Human Placenta Random Primed 2
Human Smad4_v3	5	pB27	hgx530 (HYPA; prey24113) hHYPA	Human Breast Epithelial cells RP1

Human Smad4_v3	5	pB27	hgx530 (HYPA; prey24113) hHYPA	Human CEMC7 Random Primed
Human Smad5_v2	6	pB27	prey69489 (ZNF8) hZNF8	Human Placenta Random Primed 2
Human Smad5_v2	6	pB27	prey109486 (PTPN12 PTPG1 PTP PEST) hPTPN12	Human Placenta Random Primed 2
Smad9a isoform_ v1	7	pB27	prey156850 (ZNF8 HF.18) hZNF8	Human Placenta Random Primed 3
Smad9a isoform_ v1	7	pB27	prey34730 (LMO4; prey34731) hLMO4	Human Placenta Random Primed 3
Human Smad9_v3	8	pB27	prey34730 (LMO4; prey34731) hLMO4	Human Placenta Random Primed 2
Human SMURF2_v 2	9	pB27	prey007779 - Human LAPTM5	Human Placenta Random Primed 2
Human SMURF2_v 2	9	pB27	hgx596 (LAPTm5; prey7779) hLAPTm5	Human Thymocytes Random Primed 1
Human SMURF2_v 2	9	pB27	hgx596 (LAPTm5; prey7779) hLAPTm5	Human CEMC7 Random Primed
Human SMURF2_v 2	9	pB27	prey007748 - Human RNF11	Human Placenta Random Primed 2
Human SMURF2_v 2	9	pB27	hgx555 (prey7748) hCGI 123protein RNF 11 hRNF11	Human Thymocytes Random Primed 1
Human SMURF2_v 2	9	pB27	hgx555 (prey7748) hCGI 123protein RNF 11 hRNF11	Human CEMC7 Random Primed
Human SMURF2_v 2	9	pB27	hgx555 (prey7748) hCGI 123protein RNF 11 hRNF11	Human Breast Epithelial cells RP1
Human SARA_v5	10	pB27	prey7748 (RNF11 SID1669 CGI 123 DKFZP564O033; prey7749) hRNF11 hCGI 123protein	Human Placenta Random Primed 2
Human SARA_v5	10	pB27	hgx555 (prey7748) hCGI 123protein RNF 11 hRNF11	Human Breast Epithelial cells RP1

Human SARA_v5	10	pB27	prey67613 (PPP1CA PPP1A; prey67618) hPPP1CA hserine/threonine specific proteinphosphatase	Human Placenta Random Primed 2
Human SARA_v5	10	pB27	hgx591 (PPP1CA PPP1A; prey67613) hPPP1CA hserine/threonine specific proteinphosphatase	Human Breast Epithelial cells RP1
Human SARA_v5	10	pB27	hgx591 (PPP1CA PPP1A; prey67613) hPPP1CA hserine/threonine specific proteinphosphatase	Human CEMC7 Random Primed
Human SARA_v5	10	pB27	hgx591 (PPP1CA PPP1A; prey67613) hPPP1CA hserine/threonine specific proteinphosphatase	Human Thymocytes Random Primed 1
Human SARA_v5	20	pB27	prey27181 (FLJ20037; prey27182) hFLJ20037	Human Placenta Random Primed 2
Human SARA_v5	19	pB27	hgx594 (FLJ20037; prey27181) hFLJ20037	Human CEMC7 Random Primed
Human SNIP1_F1	11	pB27	prey027803 - Human HIPK3	Human Placenta Random Primed 2
Human SNIP1_F2	12	pB27	prey027803 - Human HIPK3	Human Placenta Random Primed 2
Human hSnoN_v2	13	pB27	hgx40 (HIPK3 PKY YAK1 DYRK6) hHIPK3	Human Placenta Random Primed 2

Table 3 : SID®

1: Bait name	2: Bait nucleic acid SEQ ID No.	3: Prey name	4: SID nucleic acid ID No.	5: SID nucleic acid sequence	6: SID amino acid ID No.	7: SID amino acid sequence
Human Smad1_v2	1	prey00035 6 Human ZNF8	27	GAGGATTCATACGGGAGAAAGACCTTATATGTGCAAGGAGTGTGGGAAAGCCTTC AGCCAGAACTCCTCCCTCGTCCAGCATGAGCGCATCCACACTGGAGACAAGCCCT ACAAGTGTGCCGAATGTGGGAAGTCTTTCTGCCATAGTACACACCTTACCGTCCA TCGGAGGATTCACACTGGGGAGAAGCCCTATGAGTGTCAGGACTGTGGGAGGGCC TTCAACCAGAACTCCTCCCTGGGGCGGCACAAGAGGACACACACTGGGGAGAAGC CATACACCTGCAGTGTGTGTGGGAAATCCTTCTCTCGGACCACTTGCCTTTTCCT GCACCTGAGAACTCACACCGAGGAGAGGCCCTACGAGTGTAACCACTGCGGGAAG GGCTTCAGGCACAGCTCATCCCTGGCCCAGCACCAGCGGAAGCACGCGGGGGAGA AGCCCTTTGAGTGCCGCCAGAGGCTGATCTTTGAGCAGACGCC	65	RIHTGERPYMCKECGKAFSQNS SLVQHERIHTGDKPYKCAECGK SFCHSTHLTVHRRIHTGEKPYE CQDCGRAFNQNSSLGRHKRTHT GEKPYTCSVCGKSFSRTTCLFL HLRTHTEERPYECNHCGKGFRH SSSLAQHQRKHAGEKPFECRQR LIFEQT
Human Smad1_v2	1	hgx554	28	CTCCCTGGGGCGGCACAAGAGGACACACACTGGGGAGAAGCCATACACCTGCAGT GTGTGTGGGAAATCCTTCTCTCGGACCACTTGCCTTTTCCTGCACCTGAGAACTC ACACCGAGGAGAGGCCCTACGAGTGTAACCACTGCGGGAAGGGCTTCAGGCACAG CTCATCCCTGGCCCAGCACCAGCGGAAGCACGCGGGGGAGAAGCCCTTTGAGTGC CGCCAGAGGCTGATCTTTGAGCAGACGCCAGCTCTCACAAAGCATGAATGGACAG AAGCCCTGGGCTGTGACCCACCTTTGAGTCAAGATGAGAGGACTCACCGAAGCGA CAGACCCTTCAAATGTAATCAGTGTGGGAAGTGTTTCATTCAGAGCTCTCACCTC ATCCGGCACCAG	66	SLGRHKRTHTGEKPYTCSVCGK SFSRTTCLFLHLRTHTEERPYE CNHCGKGFRHSSSLAQHQRKHA GEKPFECRQRLIFEQTPALTKH EWTEALGCDPPLSQDERTHRSD RPFKCNQCGKCFIQSSHLIRHQ
Human Smad1_v2	1	hgx554	29	CCAGAACTCCTCCCTGGGGCGGCACAAGAGGACACACACTGGGGAGAAGCCATAC ACCTGCAGTGTGTGTGGGAAATCCTTCTCTCGGACCACTTGCCTTTTCCTGCACC TGAGAACTCACACCGAGGAGAGGCCCTACGAGTGTAACCACTGCGGGAAGGGCTT CAGGCACAGCTCATCCCTGGCCCAGCACCAGCGGAAGCACGCGGGGGAGAAGCCC TTTGAGTGCCGCCAGAGGCTGATCTTTGAGCAGACGCCAGCTCTCACAAAGCATG AATGGACAGAAGCCCTGGGCTGTGACCCACCTTTGAGTCAAGATGAGAGGACTCA CCGAAGCGACAGACCCTTCAAATGTAATCAGTGTGGGAAGTGTTTCATTCAGAGC TC	67	QNSSLGRHKRTHTGEKPYTCSV CGKSFSRTTCLFLHLRTHTEER PYECNHCGKGFRHSSSLAQHQR KHAGEKPFECRQRLIFEQTPAL TKHEWTEALGCDPPLSQDERTH RSDRPFKCNQCGKCFIQS
Human Smad1_v2	1	prey15685 0	30	CCTTACCGTCCATCGGAGGATTCACACTGGGGAGAAGCCCTATGAGTGTCAGGAC TGTGGGAGGGCCTTCAACCAGAACTCCTCCCTGGGGCGGCACAAGAGGACACACA CTGGGGAGAAGCCATACACCTGCAGTGTGTGTGGGAAATCCTTCTCTCGGACCAC TTGCCTTTTCCTGCACCTGAGAACTCACACCGAGGAGAGGCCCTACGAGTGTAAC	68	LTVHRRIHTGEKPYECQDCGRA FNQNSSLGRHKRTHTGEKPYTC SVCGKSFSRTTCLFLHLRTHTE ERPYECNHCGKGFRHSSSLAQH

				CACTGCGGGAAGGGCTTCAGGCACAGCTCATCCCTGGCCCAGCACCAGCGGAAGC ACGCGGGGGAGAAGCCCTTTGAGTGCCGCCAGAGGCTGATCTTTGAGCAGACGCC AGCTCTCACAAAGCATGAATGGACAGAAGCCCTGGG		QRKHAGEKPFECRQRLIFEQTP ALTKHEWTEAL
Human Smad1_v1	2	prey6948 9	31	TACGGAACTCACAAAAAGCCAGGTGCAGGACAAACCCTACAAATGTACTGACTGT GGGAAGTCGTTTAACCATAACGCACACCTCACCGTGCACAAGAGGATTCATACGG GAGAAAGACCTTATATGTGCAAGGAGTGTGGGAAAGCCTTCAGCCAGAACTCCTC CCTCGTCCAGCATGAGCGCATCCACACTGGAGACAAGCCCTACAAGTGTGCCGAA TGTGGGAAGTCTTTCTGCCATAGTACACACCTTACCGTCCATCGGAGGATTCACA CTGGGGAGAAGCCCTATGAGTGTCAGGACTGTGGGAGGGCCTTCAACCAGAACTC CTCCCTGGGGCGGCACAAGAGGACACACACTGGGGAGAAGCCATACACCTGCAGT GTGTGTGGGAAATCCTTCTCTCGGACCACTTGCCTTTTCCTGCACCTGAGAACTC ACACCGAGGAGAGGCCCTACGAGTGTAACCACTGCGGGAAGGGCTTCAGGCACAG CTCATCCCTGGCCCAGCACCAGCGGAAGCACGCGGGGGAGAAGCCCTTTGAGTGC CGCCAGAGGCTGATCTTTGAGCAGACGCC	69	TELTKSQVQDKPYKCTDCGKSF NHNAHLTVHKRIHTGERPYMCK ECGKAFSQNSSLVQHERIHTGD KPYKCAECGKSFCHSTHLTVHR RIHTGEKPYECQDCGRAFNQNS SLGRHKRTHTGEKPYTCSVCGK SFSRTTCLFLHLRTHTEERPYE CNHCGKGFRHSSSLAQHQRKHA GEKPFECRQRLIFEQT
Human Smad1_v1	2	hgx554	32	ATGGACCCCGAGGACGAAGGGGTAGCGGGAGTGATGTCTGTGGGGCCGCCGGCGG CCCGGCTTCAGGAACCAGTGACCTTCCGGGATGTGGCTGTGGACTTTACCCAGGA GGAATGGGGGCAGCTGGACCCTACCCAGAGGATCCTCTACCGTGACGTGATGCTG GAGACCTTTGGTCACCTGCTCTCCATAGGTCCTGAGCTTCCGAAGCCTGAAGTCA TCTCCCAGCTGGAGCAAGGGACCGAGCTATGGGTGGCTGAGAGAGGAACCACCCA GGGCTGCCATCCAGCCTGGGAGCCTCGATCTGAAAGCCAAGCATCACGCAAGGAA GAGGGCCTGCCTGAAGAGGAGCCATCCCATGTCACGGGAAGGGAAGGATTCCCGA CAGATGCTCCTTATCCCACCACGTTAGGGAAAGACAGGGAGTGTCAGAGCCAGAG TCTGGCACTCAAGGAGCAGAATAACTTGAAGCAGTTGGAATTTGGCCTCAAGGAA GCACCAGTTCAAGATCAAGGCTACAAAACTCTCAGACTCAGGGAAAACTGCGTCC TGAGTTCAAGCCCAAATCCATTCCCAGAGATCTCTAGAGGGGAGTATTTGTATAC TTACGACTCACAGATTACAGACTCAGAACATAACTCCAGCTTAGTCAGTCAGCAG ACAGGCTCCCCAGGAAAACAGCCCGGTGAAAACAGTGACTGTCACAGAGATTCCA GTCAGGCCATTCCAATTACGGAACTCACAAAAAGCCAGGTGCAGGACAAACCCTA CAAATGTACTGACTGTGGGAAGTCGTTTAACCATAACGCACACCTCACCGTGCAC AAGAGGATTCATACGGGAGAAAGACCTTATATGTGCAAGGAGTGTGGGAAAGCCT TCAGCCAGAACTCCTCCCTCGTCCAGCATGAGCGCATCCACACTGGAGACAAGCC CTACAAGTGTGCCGAATGTGGGAAGTCTTTCTGCCATAGTACACACCTTACCGTC CATCGGAGGATTCACACTGGGGAGAAGCCCTATGAGTGTCAGGACTGTGGGAGGG CCTTCAACCAGAACTCCTCCCTGGGGCGGCACAAGAGGACACACACTGGGGAGAA GCCATACACCTGCAGTGTGTGTGGGAAATCCTTCTCTCGGACCACTTGCCTTTTC CTGCACCTGAGAACTCACACCGAGGAGAGGCCCTACGAGTGTAACCACTGCGGGA AGGGCTTCAGGCACAGCTCATCCCTGGCCCAGCACCAGCGGAAGCACGCGGGGGA GAAGCCCTTTGAGTGCCGCCAGAGGCTGATCTTTGAGCAGACGCCAGCTCTCACA AAGCATGAATGGACAGAAGCCCTGGGCTGTGACCCACCTTTGAGTCAAGATGAGA	70	MDPEDEGVAGVMSVGPPAARLQ EPVTFRDVAVDFTQEEWGQLDP TQRILYRDVMLETFGHLLSIGP ELPKPEVISQLEQGTELWVAER GTTQGCHPAWEPRSESQASRKE EGLPEEEPSHVTGREGFPTDAP YPTTLGKDRECQSQSLALKEQN NLKQLEFGLKEAPVQDQGYKTL RLRENCVLSSSPNPFPEISRGE YLYTYDSQITDSEHNSSLVSQQ TGSPGKQPGENSDCHRDSSQAI PITELTKSQVQDKPYKCTDCGK SFNHNAHLTVHKRIHTGERPYM CKECGKAFSQNSSLVQHERIHT GDKPYKCAECGKSFCHSTHLTV HRRIHTGEKPYECQDCGRAFNQ NSSLGRHKRTHTGEKPYTCSVC GKSFSRTTCLFLHLRTHTEERP YECNHCGKGFRHSSSLAQHQRK HAGEKPFECRQRLIFEQTPALT KHEWTEALGCDPPLSQDERTHR SDRPFKCNQCGKCFIQSSHLIR HQITHTREEQPHGRSRRREQSS SRNSHLVQHQHPNSRKSSAGGA KAGQPESRALALFDIQKIMQEK

				GGACTCACCGAAGCGACAGACCCTTCAAATGTAATCAGTGTGGGAAGTGTTTCAT TCAGAGCTCTCACCTCATCCGGCACCAGATAACTCACACCAGAGAGGAGCAGCCC CATGGGCGAAGCCGGCGGCGTGAACAATCCTCGAGCAGGAACTCACACCTGGTTC AGCATCAACACCCGAACTCCAGAAAGAGCTCTGCAGGCGGAGCAAAGGCAGGGCA GCCGGAAAGCAGAGCCCTGGCTTTGTTTGACATCCAAAAAATCATGCAAGAGAAA AACCCTGTGCACGTTATTGGGGTGGAAGAGCCTTCTGTGGGTGCTTCCATGTTAT TTGACATCAGAGAATCCACATAG		NPVHVIGVEEPSVGASMLFDIR EST
Human Smad1_ v1	2	hgx554	33	GATTCATACGGGAGAAAGACCTTATATGTGCAAGGAGTGTGGGAAAGCCTTCAGC CAGAACTCCTCCCTCGTCCAGCATGAGCGCATCCACACTGGAGACAAGCCCTACA AGTGTGCCGAATGTGGGAAGTCTTTCTGCCATAGTACACACCTTACCGTCCATCG GAGGATTCACACTGGGGAGAAGCCCTATGAGTGTCAGGACTGTGGGAGGGCCTTC AACCAGAACTCCTCCCTGGGGCGGCACAAGAGGACACACACTGGGGAGAAGCCAT ACACCTGCAGTGTGTGTGGGAAATCCTTCTCTCGGACCACTTGCCTTTTCCTGCA CCTGAGAACTCACACCGAGGAGAGGCCCTACGAGTGTAACCACTGCGGGAAGGGC TTCAGGCACAGCTCATCCCTGGCCCAGCACCAGCGGAAGCACGCGGGGGAGAAGC CCTTTGAGTGCCGCCAGAGGCTGATCTTTGAGCAGA	71	IHTGERPYMCKECGKAFSQNSS LVQHERIHTGDKPYKCAECGKS FCHSTHLTVHRRIHTGEKPYEC QDCGRAFNQNSSLGRHKRTHTG EKPYTCSVCGKSFSRTTCLFLH LRTHTEERPYECNHCGKGFRHS SSLAQHQRKHAGEKPFECRQRL IFEQ
Human Smad1_ v1	2	hgx554	34	CTACAAGTGTGCCGAATGTGGGAAGTCTTTCTGCCATAGTACACACCTTACCGTC CATCGGAGGATTCACACTGGGGAGAAGCCCTATGAGTGTCAGGACTGTGGGAGGG CCTTCAACCAGAACTCCTCCCTGGGGCGGCACAAGAGGACACACACTGGGGAGAA GCCATACACCTGCAGTGTGTGTGGGAAATCCTTCTCTCGGACCACTTGCCTTTTC CTGCACCTGAGAACTCACACCGAGGAGAGGCCCTACGAGTGTAACCACTGCGGGA AGGGCTTCAGGCACAGCTCATCCCTGGCCCAGCACCAGCGGAAGCACGCGGGGGA GAAGCCCTTTGAGTGCCGCCAGAGGCTGATCTTTGAGCAGACGCCAGCTCTCACA AAGCATGAATGGACAGAAGCCCTGGGCTGTGACCCACCTTTGAGTCAAGATGAGA GGACTCACCGAAGCGACAGACCCTTCAAATGTAATCAGTGTGGGAAGTGTTTCAT TCAGAGCTC	72	YKCAECGKSFCHSTHLTVHRRI HTGEKPYECQDCGRAFNQNSSL GRHKRTHTGEKPYTCSVCGKSF SRTTCLFLHLRTHTEERPYECN HCGKGFRHSSSLAQHQRKHAGE KPFECRQRLIFEQTPALTKHEW TEALGCDPPLSQDERTHRSDRP FKCNQCGKCFIQS
Human Smad1_ v1	2	prey1740 9	35	CACGTCGGCCCAGGTGGCGGTGTTCCACCTGCAGGAGATAGCGGAGGACGAGCTG GCCCGCGACTGGACCAAGCGGCGCATGAAGGATGACCTTGTGCCCGAGACCGCAC GGCTCAACTACACTCGACCAGGGCTCCCCACGCTGAACCCCCAGCTGCTAGAGGC ATGGAAGAATGAAGTGAAGGAGAAAGGCCACGTCAACTGTCCCAACGACTGCTGT GAAGCCATCTACTCCAGCGTGTCCGGACTCAAGGCTCATCTCGCCAGCTGCAGTA AGGGGGCCCACCTGGCAGGGAAGTACCGCTGTCTGCTGTGTCCGAAGGAGTTCAG TTCTGAGAGTGGCGTCAAATACCACATCCTGAAGACCCACGCAGAGAACTGGTTC CGAACATCAGCAGACCCACCTCCCAAACACAGGAGCCAGGACTCATTGGTGCCCA AGAAGGAAAAGAAGAAAAATCTGGCAGGTGGAAAGAAGCGGGGCCGAAAGCCCAA GGAGCGGACCCCAGAGGAGCCTGTGGCCAAGCTGCCCCCGCGCCGGGACGACTGG CCTCCAGGATGCAGAGACAAG	73	TSAQVAVFHLQEIAEDELARDW TKRRMKDDLVPETARLNYTRPG LPTLNPQLLEAWKNEVKEKGHV NCPNDCCEAIYSSVSGLKAHLA SCSKGAHLAGKYRCLLCPKEFS SESGVKYHILKTHAENWFRTSA DPPPKHRSQDSLVPKKEKKKNL AGGKKRGRKPKERTPEEPVAKL PPRRDDWPPGCRDK
Human Smad1_	2	hgx559	36	ACCCACTGTGGCAAGACGTACCGATCCAAGGCTGGCCACGACTACCACGTGCGCT CGGAGCACACGGCCCCCCCCCCTGAGGAGCCCACAGACAAGTCCCCTGAGGCTGA	74	THCGKTYRSKAGHDYHVRSEHT APPPEEPTDKSPEAEDPLGVER

WO 03/045990 PCT/EP02/13866

v1				GGACCCGCTGGGTGTGGAGCGGACCCCAAGCGGGCGTGTCCGCCGCACGTCGGCC CAGGTGGCGGTGTTCCACCTGCAGGAGATAGCGGAGGACGAGCTGGCCCGCGACT GGACCAAGCGGCGCATGAAGGATGACCTTGTGCCCGAGACCTCACAGCTCAACTA CACTCGACCAGGGCTCCCCACGCTGAACCCCCAGCTGCTAGAGGCATGGAAGAAT GAAGTGAAGGAGAAAGGCCACGTCAACTGTCCCAACGACTGCTGTGAAGCCATCT ACTCCAGCGTGTCCGGACTCAAGGCTCATCTCGCCAGCTGCAGTAAGGGGGCCCA CCTGGCAGGGAAGTACCGCTGTCTGCTGTGTCCGAAGGAGTTCAGTTCTGAGAGT GGCGTCAAATACCACATCCTGAAGACCCACGCAGAGAACTGGTTCCGAACATCAG CAGACCCACCTCCCAAACACAGGAGCCAGGACTCATTGGTGCCCAAGAAGGAAAA GAAGAAAAATCTGGCAGGTGGAAAGAAGCGGGGCCGAAAGCCCAAGGAGCGGACC CCAGAG		TPSGRVRRTSAQVAVFHLQEIA EDELARDWTKRRMKDDLVPETS QLNYTRPGLPTLNPQLLEAWKN EVKEKGHVNCPNDCCEAIYSSV SGLKAHLASCSKGAHLAGKYRC LLCPKEFSSESGVKYHILKTHA ENWFRTSADPPPKHRSQDSLVP KKEKKKNLAGGKKRGRKPKERT PE
Human Smad1_ v3	3	prey106053	37	CACGTCGGCCCAGGTGGCGGTGTTCCACCTGCAGGAGATAGCGGAGGACGAGCTG GCCCGCGACTGGACCAAGCGGCGCATGAAGGATGACCTTGTGCCCGAGACCTCAC AGCTCAACTACACTCGACCAGGGCTCCCCACGCTGAACCCCCAGCTGCTAGAGGC ATGGAAGAATGAAGTGAAGGAGAAAGGCCACGTCAACTGTCCCAACGACTGCTGT GAAGCCATCTACTCCAGCGTGTCCGGACTCAAGGCTCATCTCGCCAGCTGCAGTA AGGGGGCCCACCTGGCAGGGAAGTACCGCTGTCTGCTGTGTCCGAAGGAGTTCAG TTCTGAGAGTGGCGTCAAATACCACATCCTGAAGACCCACGCAGAGAACTGGTTC CGAACATCA	75	TSAQVAVFHLQEIAEDELARDW TKRRMKDDLVPETSQLNYTRPG LPTLNPQLLEAWKNEVKEKGHV NCPNDCCEAIYSSVSGLKAHLA SCSKGAHLAGKYRCLLCPKEFS SESGVKYHILKTHAENWFRTS
Human Smad4_ v1	4	prey69401	38	ATGCATGGTAGAAAGGATGATGCACAAAAGCAGCCTGTTAAAAATCAGCTTGGAT TAAACCCGCAGTCACATCTACCAGAACTGCAGCTATTTCAAGCTGAAGGGAAAAT ATATAAATATGATCACATGGAAAAATCCGTCAACAGTAGTTCCTTAGTTTCCCCA CCCCAACGTATTTCTTCTACTGTCAAAACCCACATTTCTCATACATATGAATGTA ATTTTGTGGATTCATTATTCACACAAAAAGAGAAAG	76	MHGRKDDAQKQPVKNQLGLNPQ SHLPELQLFQAEGKIYKYDHME KSVNSSSLVSPPQRISSTVKTH ISHTYECNFVDSLFTQKEK
Human Smad4_ v3	5	prey024113 Human HYPA	39	CAATGAGCACCATGGCTGCTGCCGAAGCAGCAGCTGCTGTTGTTGCAGCAGCAGC AGCGGCAGCAGCAGCAGCAGCTGCAGCCAATGCTAATGCTTCCACTTCTGCTTCT AATACTGTCAGTGGAACTGTTCCAGTTGTTCCTGAGCCTGAAGTTACTTCCATTG TTGCTACTGTTGTAGATAATGAGAATACAGTAACTATTTCAACTGAGGAACAAGC ACAACTTACTAGTACCCCTGCTATTCAGGATCAAAGTGTGGAAGTATCCAGTAAT ACTGGAGAAGAAACATCTAAGCAAGAAACTGTAG	77	MSTMAAAEAAAAVVAAAAAAAA AAAAANANASTSASNTVSGTVP VVPEPEVTSIVATVVDNENTVT ISTEEQAQLTSTPAIQDQSVEV SSNTGEETSKQETV
Human Smad4_ v3	5	hgx530	40	AGAAGAGTGCACCACAACATCAACAGCCCCAGTCCCTACAACAGAAATTCCGACC ACAATGAGCACCATGGCTGCTGCCGAAGCAGCAGCTGCTGTTGTTGCAGCAGCAG CAGCGGCAGCAGCAGCAGCAGCTGCAGCCAATGCTAATGCTTCCACTTCTGCTTC TAATACTGTCAGTGGAACTGTTCCAGTTGTTCCTGAGCCTGAAGTTACTTCCATT GTTGCTACTGTTGTAGATAATGAGAATACAGTAACTATTTCAACTGAGGAACAAG CACAACTTACTAGTACCCCTGCTATTCAGGATCAAAGTGTGGAAGTATCCAGTAA TACTGGAGAAGAAACATCTAAGCAAGAAACTGTAGCTGATTTTACTCCCAAAAAA GAAGAGGAGGAGAGCCAACCAGCAAAGAAAACATACACTTGGAATACAAAGGAAG AGGCAAAGCAAGCTTTTAAAGAATTATTGAAAGAAAAGCG	78	EECTTTSTAPVPTTEIPTTMST MAAAEAAAAVVAAAAAAAAAAA AANANASTSASNTVSGTVPVVP EPEVTSIVATVVDNENTVTIST EEQAQLTSTPAIQDQSVEVSSN TGEETSKQETVADFTPKKEEEE SQPAKKTYTWNTKEEAKQAFKE LLKEK

Human Smad4_ v3	5	hgx530	41	CTGCCGAAGCAGCAGCTGCTGTTGTTGCAGCAGCAGCAGCGGCAGCAGCAGCAGC AGCTGCAGCCAATGCTAATGCTTCCACTTCTGCTTCTAATACTGTCAGTGGAACT GTTCCAGTTGTTCCTGAGCCTGAAGTTACTTCCATTGTTGCTACTGTTGTAGATA ATGAGAATACAGTAACTATTTCAACTGAGGAACAAGCACAACTTACTAGTACCCC TGCTATTCAGGATCAAAGTGTGGAAGTATCCAGTAATACTGGAGAAGAAACATCT AAGCAAGAAACTGTAGCTGATTTTACTCCCAAAAAAGAAGAGGAGGAGAGCCAAC	79	AEAAAAVVAAAAAAAAAAAAAN ANASTSASNTVSGTVPVVPEPE VTSIVATVVDNENTVTISTEEQ AQLTSTPAIQDQSVEVSSNTGE ETSKQETVADFTPKKEEEESQ
Human Smad5_ v2	6	prey6948 9	42	GAGGATTCATACGGGAGAAAGACCTTATATGTGCAAGGAGTGTGGGAAAGCCTTC AGCCAGAACTCCTCCCTCGTCCAGCATGAGCGCATCCACACTGGAGACAAGCCCT ACAAGTGTGCCGAATGTGGGAAGTCTTTCTGCCATAGTACACACCTTACCGTCCA TCGGAGGATTCACACTGGGGAGAAGCCCTATGAGTGTCAGGACTGTGGGAGGGCC TTCAACCAGAACTCCTCCCTGGGGCGGCACAAGAGGACACACACTGGGGAGAAGC CATACACCTGCAGTGTGTGTGGGAAATCCTTCTCTCGGACCACTTGCCTTTTCCT GCACCTGAGAACTCACACCGAGGAGAGGCCCTACGAGTGTAACCACTGCGGGAAG GGCTTCAGGCACAGCTCATCCCTGGCCCAGCACCAGCGGAAGCACGCGGGGGAGA AGCCCTTTGAGTGCCGCCAGAGGCTGATCTTTGAGCAGACGCC	80	RIHTGERPYMCKECGKAFSQNS SLVQHERIHTGDKPYKCAECGK SFCHSTHLTVHRRIHTGEKPYE CQDCGRAFNQNSSLGRHKRTHT GEKPYTCSVCGKSFSRTTCLFL HLRTHTEERPYECNHCGKGFRH SSSLAQHQRKHAGEKPFECRQR LIFEQT
Human Smad5_ v2	6	prey1094 86	43	CCAAAAGCATATGTAGCAACTCAAGGACCTTTAGCAAATACAGTAATAGATTTTT GGAGGATGATATGGGAGTATAATGTTGTGATCATTGTAATGGCCTGCCGAGAATT TGAGATGGGAAGGAAAAAATGTGAGCGCTATTGGCCTTTGTATGGAGAAGACCCC ATAACGTTTGCACCATTTAAAATTTCTTGTGAGGATGAACAAGCAAGAACAGACT ACTTCATCAGGACACTCTTACTTGAATTTCAAAATGAATCTCGTAGGCTGTATCA GTTTCATTATGTGAACTGGCCAGACCATGATGTTCCTTCATCATTTGATTCTATT CTGGACATGATAAGCTTAATGAGGAAATATCAAGAACATGAAGATGTTCCTATTT GTATTCATTGCAGTGCAGGCTGTGGAAGAACAGGTGCCATTTGTGCCATAGATTA TACGTGGAATTTACTAAAAGCTGGGAAAATACCAGAGGAATTTAATGTATTTAAT TTAATACAAGAAATGAGAACACAAAGGCATTCTGCAGTACAAACAAAGGAGCAAT ATGAACTTGTTCATAGAGCTATTGCCCAACTGTTTGAAAAACAGCTACAACTATA TGAAATTCATGGAGCTCAGAAAATTGCTGATGGAGTGAATGAAATTAACACTGAA AACATGATCAGCTCCATAGAGCCTGAAAAACAAGATTCTCCTCCTCCAAAACC	81	PKAYVATQGPLANTVIDFWRMI WEYNVVIIVMACREFEMGRKKC ERYWPLYGEDPITFAPFKISCE DEQARTDYFIRTLLLEFQNESR RLYQFHYVNWPDHDVPSSFDSI LDMISLMRKYQEHEDVPICIHC SAGCGRTGAICAIDYTWNLLKA GKIPEEFNVFNLIQEMRTQRHS AVQTKEQYELVHRAIAQLFEKQ LQLYEIHGAQKIADGVNEINTE NMISSIEPEKQDSPPPK
Smad9a isofor m_v1	7	prey3473 0	44	CTCGCAGCCGCCCCCGGTGACGGCCGGCTCCCTCTCCTGGAAGCGGTGCGCAGGC TGCGGGGGCAAGATTGCGGACCGCTTTCTGCTCTATGCCATGGACAGCTATTGGC ACAGCCGGTGCCTCAAGTGCTCCTGCTGCCAGGCGCAGCTGGGCGACATCGGCAC GTCCTGTTACACCAAAAGTGGCATGATCCTTTGCAGAAATGACTACATTAGGTTA TTTGGAAATAGCGGTGCTTGCAGCGCTTGCGGACAGTCGATTCCTGCGAGTGAAC TCGTCATGAGGGCGCAAGGCAATGTGTATCATCTTAAGTGTTTTACATGCTCTAC CTGCCGGAATCGCCTGGTCCCGGGAGATCGGTTTCACTACATCAATGGCAGTTTA TTTTGTGAACATGATAGACCTACAGCTCTCATCAATGGCCATTTGAATTCACTTC AGAGCAATCCACTACTGCCAGACCAGAAGGTCTGCTAA	82	SQPPPVTAGSLSWKRCAGCGGK IADRFLLYAMDSYWHSRCLKCS CCQAQLGDIGTSCYTKSGMILC RNDYIRLFGNSGACSACGQSIP ASELVMRAQGNVYHLKCFTCST CRNRLVPGDRFHYINGSLFCEH DRPTALINGHLNSLQSNPLLPD QKVC
Smad9a isofor	7	Prey1568 50	45	CTCAGAACATAACTCCAGCTTAGTCAGTCAGCAGACAGGCTCCCCAGGAAAACAG CCCGGTGAAAACAGTGACTGTCACAGAGATTCCAGTCAGGCCATTCCAATTACGG	83	SEHNSSLVSQQTGSPGKQPGEN SDCHRDSSQAIPITELTKSQVQ

m_v1				AACTCACAAAAAGCCAGGTGCAGGACAAACCCTACAAATGTACTGACTGTGGGAA GTCGTTTAACCATAACGCACACCTCACCGTGCACAAGAGGATTCATACGGGAGAA AGACCTTATATGTGCAAGGAGTGTGGGAAAGCCTTCAGCCAGAACTCCTCCCTCG TCCAGCATGAGCGCATCCACACTGGAGACAAGCCCTACAAGTGTGCCGAATGTGG GAAGTCTTTCTGCCATAGTACACACCTTACCGTCCATCGGAGGATTCACACTGGG GAGAAGCCCTATGAGTGTCAGGACTGTGGGAGGGCCTTCAACCAGAACTCCTCCC TGGGGCGGCACAAGAGGACACACACTGGGGAGAAGCCATACACCTGCAGTGTGTG TGGGAAATCCTTCTCTCGGACCACTTGCCTTTTCCTGCACCTGAGAACTCACACC GAGGAGAGGCCCTACGAGTGTAACCACTGCGGGAAGGGCTTCAGGCACAGCTCAT CCCTGGCCCAGCACCAGCGGAAGCAC		DKPYKCTDCGKSFNHNAHLTVH KRIHTGERPYMCKECGKAFSQN SSLVQHERIHTGDKPYKCAECG KSFCHSTHLTVHRRIHTGEKPY ECQDCGRAFNQNSSLGRHKRTH TGEKPYTCSVCGKSFSRTTCLF LHLRTHTEERPYECNHCGKGFR HSSSLAQHQRKH
Human Smad9_ v3	8	prey3473 0	46	CTCGCAGCCGCCCCCGGTGACGGCCGGCTCCCTCTCCTGGAAGCGGTGCGCAGGC TGCGGGGGCAAGATTGCGGACCGCTTTCTGCTCTATGCCATGGACAGCTATTGGC ACAGCCGGTGCCTCAAGTGCTCCTGCTGCCAGGCGCAGCTGGGCGACATCGGCAC GTCCTGTTACACCAAAAGTGGCATGATCCTTTGCAGAAATGACTACATTAGGTTA TTTGGAAATAGCGGTGCTTGCAGCGCTTGCGGACAGTCGATTCCTGCGAGTGAAC TCGTCATGAGGGCGCAAGGCAATGTGTATCATCTTAAGTGTTTTACATGCTCTAC CTGCCGGAATCGCCTGGTCCCGGG	84	SQPPPVTAGSLSWKRCAGCGGK IADRFLLYAMDSYWHSRCLKCS CCQAQLGDIGTSCYTKSGMILC RNDYIRLFGNSGACSACGQSIP ASELVMRAQGNVYHLKCFTCST CRNRLVP
Human SMURF2 _v2	9	prey0077 79LAPTM5	47	GAAGACCCCAGAGGGGGGCCCAGCACCACCCCCATACTCAGAGGTGTGA	85	KTPEGGPAPPPYSEV*
Human SMURF2 _v2	9	hgx596	48	GACCCCAGAGGGGGGCCCAGCACCACCCCCATACTCAGAGGTGTGA	86	TPEGGPAPPPYSEV
Human SMURF2 _v2	9	hgx596	49	GGGGGGCCCAGCACCACCCCCATACTCAGAGGTGTGA	87	GGPAPPPYSEV
Human SMURF2 _v2	9	prey0077 48 Human RNF11	50	GCCGGATCAGGAGCCGCCGCCGCCATATCAGGAACAAGTTCCAGTTCCAGTCTAC CACCCAACACCTAGCCAGACTCGGCTAGCAACTCAGCTGACTGAAGAGGAACAAA TTAGGATAGCTCAAAGAATAGGTCTTATACAACATCTGCCTAAAGGAGTTTATGA CCCTGGAAGAGATGGATCAGAAAAAAAGATCCGGGAGTGTGTGATCTGTATGATG GACTTTGTTT	88	PDQEPPPPYQEQVPVPVYHPTP SQTRLATQLTEEEQIRIAQRIG LIQHLPKGVYDPGRDGSEKKIR ECVICMMDFV
Human SMURF2 _v2	9	hgx555	51	ATGGGGAACTGCCTCAAATCCCCCACCTCGGATGACATCTCCCTGCTTCACGAGT CTCAGTCCGACCGGGCTAGCTTTGGCGAGGGGACGGAGCCGGATCAGGAGCCGCC GCCGCCATATCAGGAACAAGTTCCAGTTCCAGTCTACCACCCAACACCTAGCCAG ACTCGGCTAGCAACTCAGCTGACTGAAGAGGAACAAATTAGGATAGCTCAAAGAA TAGGTCTTATACAACATCTGCCTAAAGGAGTTTATGACCCTGGAAGAGATGGATC AGAAAAAAAGATCCG	89	MGNCLKSPTSDDISLLHESQSD RASFGEGTEPDQEPPPPYQEQV PVPVYHPTPSQTRLATQLTEEE QIRIAQRIGLIQHLPKGVYDPG RDGSEKKI
Human SMURF2	9	hgx555	52	ATGGGGAACTGCCTCAAATCCCCCACCTCGGATGACATCTCCCTGCTTCACGAGT CTCAGTCCGACCGGGCTAGCTTTGGCGAGGGGACGGAGCCGGATCAGGAGCCGCC	90	MGNCLKSPTSDDISLLHESQSD RASFGEGTEPDQEPPPPYQEQV

_v2				GCCGCCATATCAGGAACAAGTTCCAGTTCCAGTCTACCACCCAACACCTAGCCAG ACTCGGCTAGCAACTCAGCTGACTGAAGAGGAACAAATTAGGATAGCTCAAAGAA TAGGTCTTATACAACATCTGCCTAAAGGAG		PVPVYHPTPSQTRLATQLTEEE QIRIAQRIGLIQHLPKG
Human SMURF2 _v2	9	hgx555	53	ATGGGGAACTGCCTCAAATCCCCCACCTCGGATGACATCTCCCTGCTTCACGAGT CTCAGTCCGACCGGGCTAGCTTTGGCGAGGGGACGGAGCCGGATCAGGAGCCGCC GCCGCCATATCAGGAACAAGTTCCAGTTCCAGTCTACCACCCAACACCTAGCCAG ACTCGGCTAGCAACTCAGCTGACTGAAGAGGAACAAATTAGGATAGCTCAAAGAA TAGGTCTTATACAACATCTGCCTAAAGGAGTTTATGACCCTGGAAGAGATGGATC AGAAAAAAAGATCCGGGAGTGTGTGATCTGTATGATGGACTTTGTTTATGGGGAC CCAATTCGATTTCTGCCGTGCATGCACATCTATCACCTGGACTGTATAGATGACT GGTTGATGAGATCCTTCACGTGCCCCTCCTGCATGGAGCCAGTTGATGCAGC	91	MGNCLKSPTSDDISLLHESQSD RASFGEGTEPDQEPPPPYQEQV PVPVYHPTPSQTRLATQLTEEE QIRIAQRIGLIQHLPKGVYDPG RDGSEKKIRECVICMMDFVYGD PIRFLPCMHIYHLDCIDDWLMR SFTCPSCMEPVDA
Human SARA_v 5	10	prey7748	54	GCTGACTGAAGAGGAACAAATTAGGATAGCTCAAAGAATAGGTCTTATACAACAT CTGCCTAAAGGAGTTTATGACCCTGGAAGAGATGGATCAGAAAAAAAGATCCGGG AGTGTGTGATCTGTATGATGGACTTTGTTTATGGGGACCCAATTCGATTTCTGCC GTGCATGCACATCTATCACCTGGACTGTATAGATGACTGGTTGATGAGATCCTTC ACGTGCCCCTCCTGCATGGAGCCAGTTGATGCAGCACTGCTTTCATCCTATGAGA CTAATTGA	92	LTEEEQIRIAQRIGLIQHLPKG VYDPGRDGSEKKIRECVICMMD FVYGDPIRFLPCMHIYHLDCID DWLMRSFTCPSCMEPVDAALLS SYETN
Human SARA_v 5	10	hgx555	55	CCAGACTCGGCTAGCAACTCAGCTGACTGAAGAGGAACAAATTAGGATAGCTCAA AGAATAGGTCTTATACAACATCTGCCTAAAGGAGTTTATGACCCTGGAAGAGATG GATCAGAAAAAAAGATCCGGGAGTGTGTGATCTGTATGATGGACTTTGTTTATGG GGACCCAATTCGATTTCTGCCGTGCATGCACATCTATCACCTGGACTGTATAGAT GACTGGTTGATGAGATCCTTCACGTGCCCCTCCTGCATGGAGCCAGTTGATGCAG CACTGCTTTCATCCTATGAGACTAATTGA	93	QTRLATQLTEEEQIRIAQRIGL IQHLPKGVYDPGRDGSEKKIRE CVICMMDFVYGDPIRFLPCMHI YHLDCIDDWLMRSFTCPSCMEP VDAALLSSYETN
Human SARA_v 5	10	prey6761 3	56	GCTGACAGAGAACGAGATCCGCGGTCTGTGCCTGAAATCCCGGGAGATTTTTCTG AGCCAGCCCATTCTTCTGGAGCTGGAGGCACCCCTCAAGATCTGCGGTGACATAC ACGGCCAGTACTACGACCTTCTGCGACTATTTGAGTATGGCGGTTTCCCTCCCGA GAGCAACTACCTCTTTCTGGGGGACTATGTGGACAGGGGCAAGCAGTCCTTGGAG ACCATCTGCCTGCTGCTGGCCTATAAGATCAAGTACCCCGAGAACTTCTTCCTGC TCCGTGGGAACCACGAGTGTGCCAGCATCAACCGCATCTATGGTTTCTACGATGA GTGCAAGAGACGCTACAACATCAAACTGTGGAAAACCTTCACTGACTGCTTCAAC TGCCTGCCCATCGCGGCCATAGTGGACGAAAAGATCTTCTGCTGCCACGGAGGCC TGTCCCCGGACCTGCAGTCTATGGAGCAGATTCGGCGGATCATGCGGCCCACAGA TGTGCCTGACCAGGGCCTGCTGTGTGACCTGCTGTGGTCTGACCCTGACAAGGAC GTGCAGGGCTGGGGCGAGAACGACCGTGGCGTCTCTTTTACCTTTGGAGCCGAGG TGGTGGCCAAGTTCCTCCACAAGCACGACTTGGACCTCATCTGCCGAGCACACCA GGTGGTAGAAGACGGCTACGAGTTCTTTGCCAAGCGGCAGCTGGTGACACTTTTC TCAGCTCCCAACTACTGTGGCGAGTTTGACAATGCTGGCGCCATGATGAGTGTGG ACGAGACCCTCATGTGCTCTTTCCAGATC	94	LTENEIRGLCLKSREIFLSQPI LLELEAPLKICGDIHGQYYDLL RLFEYGGFPPESNYLFLGDYVD RGKQSLETICLLLAYKIKYPEN FFLLRGNHECASINRIYGFYDE CKRRYNIKLWKTFTDCFNCLPI AAIVDEKIFCCHGGLSPDLQSM EQIRRIMRPTDVPDQGLLCDLL WSDPDKDVQGWGENDRGVSFTF GAEVVAKFLHKHDLDLICRAHQ VVEDGYEFFAKRQLVTLFSAPN YCGEFDNAGAMMSVDETLMCSF QI
Human	10	hgx591	57	GCGGCCTGGCAAGAATGTACAGCTGACAGAGAACGAGATCCGCGGTCTGTGCCTG	95	RPGKNVQLTENEIRGLCLKSRE

WO 03/045990 PCT/EP02/13866

SARA_v5				AAATCCCGGGAGATTTTTCTGAGCCAGCCCATTCTTCTGGAGCTGGAGGCACCCC TCAAGATCTGCGGTGACATACACGGCCAGTACTACGACCTTCTGCGACTATTTGA GTATGGCGGTTTCCCTCCCGAGAGCAACTACCTCTTTCTGGGGGACTATGTGGAC AGGGGCAAGCAGTCCTTGGAGACCATCTGCCTGCTGCTGGCCTATAAGATCAAGT ACCCCGAGAACTTCTTCCTGCTCCGTGGGAACCACGAGTGTGCCAGCATCAACCG CATCTATGGTTTCTACGATGAGTGCAAGAGACGCTACAACATCAAACTGTGGAAA ACCTTCACTGACTGCTTCAACTGCCTGCCCATCGCGGCCATAGTGGACGAAAAGA TCTTCTGCTGCCACGGAGGCCTGTCCCCGGACCTGCAGTCTATGGAGCAGATTCG GCGGATCATGCGGCCCACAGATGTGCCTGACCAGGGCCTGCTGTGTGACCTGCTG TGGTCTGACCCTGACAAGGACGTGCAGGGCTGGGGCGAGAACGACCGTGGCGTCT CTTTTACCTTTGGAGCCGAGGTGGTGGCCAAGTTCCTCCACAAGCACGACTTGGA CCTCATCTGCCGAGCACACCAGGTGGTAGAAGACGGCTACGAGTTCTTTGCCAAG CGGCAGCTGGTGACACTTTTCTCAGCTCCCAACTACTGTGGCGAGTTTGACAATG CTGGCGCCATGATGAGTGTGGACGAGACCCTCATGTGCTCTTTCCAGATCCTCAA GCC		IFLSQPILLELEAPLKICGDIH GQYYDLLRLFEYGGFPPESNYL FLGDYVDRGKQSLETICLLLAY KIKYPENFFLLRGNHECASINR IYGFYDECKRRYNIKLWKTFTD CFNCLPIAAIVDEKIFCCHGGL SPDLQSMEQIRRIMRPTDVPDQ GLLCDLLWSDPDKDVQGWGEND RGVSFTFGAEVVAKFLHKHDLD LICRAHQVVEDGYEFFAKRQLV TLFSAPNYCGEFDNAGAMMSVD ETLMCSFQILK
Human SARA_v5	10	hgx591	58	ATGTCCGACAGCGAGAAGCTCAACCTGGACTCGATCATCGGGCGCCTGCTGGAAG TGCAGGGCTCGCGGCCTGGCAAGAATGTACAGCTGACAGAGAACGAGATCCGCGG TCTGTGCCTGAAATCCCGGGAGATTTTTCTGAGCCAGCCCATTCTTCTGGAGCTG GAGGCACCCCTCAAGATCTGCGGTGACATACACGGCCAGTACTACGACCTTCTGC GACTATTTGAGTATGGCGGTTTCCCTCCCGAGAGCAACTACCTCTTTCTGGGGGA CTATGTGGACAGGGGCAAGCAGTCCTTGGAGACCATCTGCCTGCTGCTGGCCTAT AAGATCAAGTACCCCGAGAACTTCTTCCTGCTCCGTGGGAACCACGAGTGTGCCA GCATCAACCGCATCTATGGTTTCTACGATGAGTGCAAGAGACGCTACAACATCAA ACTGTGGAAAACCTTCACTGACTGCTTCAACTGCCTGCCCATCGCGGCCATAGTG GACGAAAAGATCTTCTGCTGCCACGGAGGCCTGTCCCCGGACCTGCAGTCTATGG AGCAGATTCGGCGGATCATGCGGCCCACAGATGTGCCTGACCAGGGCCTGCTGTG TGACCTGCTGTGGTCTGACCCTGACAAGGACGTGCAGGGCTGGGGCGAGAACGAC CGTGGCGTCTCTTTTACCTTTGGAGCCGAGGTGGTGGCCAAGTTCCTCCACAAGC ACGACTTGGACCTCATCTGCCGAGCACACCAGGTGGTAGAAGACGGCTACGAGTT CTTTGCCAAGCGGCAGCTGGTGACACTTTTCTCAGCTCCCAACTACTGTGGCGAG TTTGACAATGCTGGCGCCATGATGAGTGTGGACGAGACCCTCATGTGCTCTTTCC AGATCCTCAAGCCCGCCGACAAGAACAAGGGGAAGTACGGGCAGTTCAGTGGCCT GAACCCTGGAGGCCGACCCATCACCCCACCCCGCAATTCCGCCAAAG	96	MSDSEKLNLDSIIGRLLEVQGS RPGKNVQLTENEIRGLCLKSRE IFLSQPILLELEAPLKICGDIH GQYYDLLRLFEYGGFPPESNYL FLGDYVDRGKQSLETICLLLAY KIKYPENFFLLRGNHECASINR IYGFYDECKRRYNIKLWKTFTD CFNCLPIAAIVDEKIFCCHGGL SPDLQSMEQIRRIMRPTDVPDQ GLLCDLLWSDPDKDVQGWGEND RGVSFTFGAEVVAKFLHKHDLD LICRAHQVVEDGYEFFAKRQLV TLFSAPNYCGEFDNAGAMMSVD ETLMCSFQILKPADKNKGKYGQ FSGLNPGGRPITPPRNSAK
Human SARA_v5	10	hgx591	59	GCGGCCTGGCAAGAATGTACAGCTGACAGAGAACGAGATCCGCGGTCTGTGCCTG AAATCCCGGGAGATTTTTCTGAGCCAGCCCATTCTTCTGGAGCTGGAGGCACCCC TCAAGATCTGCGGTGACATACACGGCCAGTACTACGACCTTCTGCGACTATTTGA GTATGGCGGTTTCCCTCCCGAGAGCAACTACCTCTTTCTGGGGGACTATGTGGAC AGGGGCAAGCAGTCCTTGGAGACCATCTGCCTGCTGCTGGCCTATAAGATCAAGT ACCCCGAGAACTTCTTCCTGCTCCGTGGGAACCACGAGTGTGCCAGCATCAACCG	97	RPGKNVQLTENEIRGLCLKSRE IFLSQPILLELEAPLKICGDIH GQYYDLLRLFEYGGFPPESNYL FLGDYVDRGKQSLETICLLLAY KIKYPENFFLLRGNHECASINR IYGFYDECKRRYNIKLWKTFTD

				CATCTATGGTTTCTACGATGAGTGCAAGAGACGCTACAACATCAAACTGTGGAAA ACCTTCACTGACTGCTTCAACTGCCTGCCCATCGCGGCCATAGTGGACGAAAAGA TCTTCTGCTGCCACGGAGGCCTGTCCCCGGACCTGCAGTCTATGGAGCAGATTCG GCGGATCATGCGGCCCACAGATGTGCCTGACCAGGGCCTGCTGTGTGACCTGCTG TGGTCTGACCCTGACAAGGACGTGCAGGGCTGGGGCGAGAACGACCGTGGCGTCT CTTTTACCTTTGGAGCCGAGGTGGTGGCCAAGTTCCTCCACAAGCACGACTTGGA CCTCATCTGCCGAGCACACCAGGTGGTAGAAGACGGCTACGAGTTCTTTGCCAAG CGGCAGCTGGTGACACTTTTCTCAGCTCCCAACTACTGTGGCGAGTTTGACAATG CTGGCGCCATGATGAGTGTGGACGAGACCCTCATGTGCTCTTTCCAGATCCTCAA GCCCGCCGACAAGAACAAGGGGAAGTACGGGCAGTTCAGTGGCCTGAACCCTGGA GGCCGACCCATCACCCCACCCCGCAATTCCGCCAAAGCCAAGAAATAG		CFNCLPIAAIVDEKIFCCHGGL SPDLQSMEQIRRIMRPTDVPDQ GLLCDLLWSDPDKDVQGWGEND RGVSFTFGAEVVAKFLHKHDLD LICRAHQVVEDGYEFFAKRQLV TLFSAPNYCGEFDNAGAMMSVD ETLMCSFQILKPADKNKGKYGQ FSGLNPGGRPITPPRNSAKAKK
Human SARA_v5	20	prey27181	60	GGACTATTGCGAAAGCCCTACGGCGCACTGCAATGTGCTGAACTGGGAGCAAGTG CAGCGGCTGGACGGCATCCTGAGTGAGACCATTCCGATTCACGGGCGCGGCAACT TCCCCACGCTCGAGCTGCAGCCGAGCCTGATCGTGAAGGTGGTGCGGCGGCGCCT GGCCGAGAAGCGCATTGGCGTCCGCGACGTGCGCCTCAACGGCTCGGCAGCCAGC CATGTCCTGCACCAGGACAGCGGCCTGGGCTACAAGGACCTGGACCTCATCTTCT GCGCCGACCTGCGCGGGGAAGGGGAGTTTCAGACTGTGAAGGACGTCGTGCTGGA CTGCCTGTTGGACTTCTTACCCGAGGGGGTGAACAAAGAGAAGATCACACCACTC ACGCTCAAGGAAGCTTATGTGCAGAAAATGGTTAAAGTGTGCAATGACTCTGACC GATGGAGTCTTATATCCCTGTCAAACAACAGTGGCAAAAATGTGGAACTGAAATT TGTGGATTCCCTCCGGAGGCAGTTTGAATTCAGTGTAGATTCTTTTCAAATCAAA TTAGACTCTCTTCTGCTCTTTTATGAATGTTCAGA	98	DYCESPTAHCNVLNWEQVQRLD GILSETIPIHGRGNFPTLELQP SLIVKVVRRRLAEKRIGVRDVR LNGSAASHVLHQDSGLGYKDLD LIFCADLRGEGEFQTVKDVVLD CLLDFLPEGVNKEKITPLTLKE AYVQKMVKVCNDSDRWSLISLS NNSGKNVELKFVDSLRRQFEFS VDSFQIKLDSLLLFYECS
Human SARA_v5	19	hgx594	61	ATTAGACTCTCTTCTGCTCTTTTATGAATGTTCAGAGAACCCAATGACTGAGACA TTTCACCCCACAATAATCGGGGAGAGCGTCTATGGCGATTTCCAGGAAGCCTTTG ATCACCTTTGTAACAAGATCATTGCCACCAGGAACCCAGAGGAAATCCGAGGGGG AGGCCTGCTTAAGTACTGCAACCTCTTGGTGAGGGGCTTTAGGCCCGCCTCTGAT GAAATCAAGGCCCTTCAAAGGTACATGTGTTCCAGGTTTTTCATCGACTTCTCAG ACATTGGAGAGCAGCAGAGAAAACTGGAGTCCTATTTGCAGAACCACTTTGTGGG ATTGGAAGACCGCAAGTATGAGTATCTCATGACCCTTCATGGAGTGGTAAATGAG AGCACAGTGTGCCTGATGGGACATGAAAGAAGACAGACTTTAAACCTTATCACCA TGCTGGCTATCCGGGTGTTAGCTGACCAAAATGTCATTCCTAATGTGGCTAATGT CACTTGCTATTACCAGCC	99	LDSLLLFYECSENPMTETFHPT IIGESVYGDFQEAFDHLCNKII ATRNPEEIRGGGLLKYCNLLVR GFRPASDEIKALQRYMCSRFFI DFSDIGEQQRKLESYLQNHFVG LEDRKYEYLMTLHGVVNESTVC LMGHERRQTLNLITMLAIRVLA DQNVIPNVANVTCYYQ
Human SNIP1_F1	11	prey027803 Human HIPK3	62	AAGAAATTGCTGTGAAACATCTATCAGACAGGACTCTGATTCATCAGTTTCAGAC AAACAGCGGCAAACCATCATTATTGCCGACTCCCCGAGTCCTGCAGTGAGTGTCA TCACTATCAGCAGTGACACTGATGAGGAAGAGACTTCCCAGAGACATTCACTCAG AGAATGTAAAGGTAGTCTAGATTGTGAAGCTTGCCAGAGCACTTTGAATATTGAT CGGATGTGTTCATTAAGTAGTCCTGATAGTACTCTGAGTACCAGCTCCTCAGGGC AGTCCAGCCCATCCCCCTGCAAGAGACCGAATAGTATGTCAGATGAAGAGCAAGA AAGTAGTTGTGATACGGTGGATGGCTCTCCGACATCTGACTCTTCCGGGCATGAC	100	RNCCETSIRQDSDSSVSDKQRQ TIIIADSPSPAVSVITISSDTD EEETSQRHSLRECKGSLDCEAC QSTLNIDRMCSLSSPDSTLSTS SSGQSSPSPCKRPNSMSDEEQE SSCDTVDGSPTSDSSGHDSPFA ESTFVEDTHENTELVSSADTET

WO 03/045990 PCT/EP02/13866

				AGTCCATTTGCAGAGAGCACTTTTGTGGAGGACACTCATGAAAACACAGAATTGG TATCCTCTGCTGACACAGAAACCAAGCCAGCTGTCTGTTCTGTTGTGGTGCCACC AGTGGAACTAGAAAATGGCTTAAATGCCGATGAGCATATGGCAAACACAGATTCT ATATGCCAGCCATTAATAAAAGGACGATCTGCCCCTGGAAGATTAAACCAGCCTT CTGCAGTGGGTACTCGTCAGCAAAAATTGACATCAGCATTCC		KPAVCSVVVPPVELENGLNADE HMANTDSICQPLIKGRSAPGRL NQPSAVGTRQQKLTSAF
Human SNIP1_ F2	12	prey027803 Human HIPK3	63	TGTGGGGATTGCACATGTTGTCTGGCCTCAGCCTGCCACTACCAAGAAAAATAAA CAGTGCCAGAACAGAGGTATTTTGGTAAAACTAATGGAATGGGAGCCAGGAAGAG AGGAAATAAATGCTTTCAGTTGGAGTAATTCATTACAGAATACCAATATCCCACA TTCAGCATTTATTTCTCCAAAGATAATTAATGGGAAAGATGTCGAGGAAGTAAGT TGTATAGAAACACAGGACAATCAGAACTCAGAAGGAGAGGCAAGAAATTGCTGTG AAACATCTATCAGACAGGACTCTGATTCATCAGTTTCAGACAAACAGCGGCAAAC CATCATTATTGCCGACTCCCCGAGTCCTGCAGTGAGTGTCATCACTATCAGCAGT GACACTGATGAGGAAGAGACTTCCCAGAGACATTCACTCAGAGAATGTAAAGGTA GTCTAGATTGTGAAGCTTGCCAGAGCACTTTGAATATTGATCGGATGTGTTCATT AAGTAGTCCTGATAGTACTCTGAGTACCAGCTCCTCAGGGCAGTCCAGCCCATCC C	101	VGIAHVVWPQPATTKKNKQCQN RGILVKLMEWEPGREEINAFSW SNSLQNTNIPHSAFISPKIING KDVEEVSCIETQDNQNSEGEAR NCCETSIRQDSDSSVSDKQRQT IIIADSPSPAVSVITISSDTDE EETSQRHSLRECKGSLDCEACQ STLNIDRMCSLSSPDSTLSTSS SGQSSPS
Human hSnoN_ v2	13	hgx40	64	CCCACATTCAGCATTTATTTCTCCAAAGATAATTAATGGGAAAGATGTCGAGGAA GTAAGTTGTATAGAAACACAGGACAATCAGAACTCAGAAGGAGAGGCAAGAAATT GCTGTGAAACATCTATCAGACAGGACTCTGATTCATCAGTTTCAGACAAACAGCG GCAAACCATCATTATTGCCGACTCCCCGAGTCCTGCAGTGAGTGTCATCACTATC AGCAGTGACACTGATGAGGAAGAGACTTCCCAGAGACATTCACTCAGAGAATGTA AAGGTAGTCTAGATTGTGAAGCTTGCCAGAGCACTTTGAATATTGATCGGATGTG TTCATTAAGTAGTCCTGATAGTACTCTGAGTACCAGCTCCTCAGGGCAGTCCAGC CCATCCCCCTGCAAGAGACCGAATAGTATGTCAGATGAAGAGCAAGAAAGTAGTT GTGATACGGTGGATGGCTCTCCGACATCTGACTCTTCCGGGCATGACAGTCCATT TGCAGAGAGCACTTTTGTGGAGGACACTCATGAAAACACAGAATTGGTATCCTCT GCTGACACAGAAACCAAGCCAGCTGTCTGTTCTGTTGTGGTGCCACCAGTGGAAC TAGAAAATGGCTTAAATGCCGATGAGCATATGGCAAACACAGATTCTATATGCCA GCCATTAATAAAAGGACGATCTGCCCCTGGAAGATTAAACCAGCCTTCTGCAGTG GGTACTCGTCAGCAAAAATTGACATCAGCATTCCAGCAGCAGCATTTGAACTTCA GTCAGGTTCAGCACTTTGGATCTGGGCATCAAGAGTGGAATGGAAACTTTGGGCA CAGAAGACAGCAAGCTTATATTCCTACTAGTGTTACCAGTAATCCATTCACTCTT TCTCATGGAAGTCCCAATCACACAGCAGTGCATGCCCACCTGGCTGGAAATACAC ACCTCGGAGGACAGCCTACTCTACTTCCATACCCATCATCAGCCACCC	102	PHSAFISPKIINGKDVEEVSCI ETQDNQNSEGEARNCCETSIRQ DSDSSVSDKQRQTIIIADSPSP AVSVITISSDTDEEETSQRHSL RECKGSLDCEACQSTLNIDRMC SLSSPDSTLSTSSSGQSSPSPC KRPNSMSDEEQESSCDTVDGSP TSDSSGHDSPFAESTFVEDTHE NTELVSSADTETKPAVCSVVVP PVELENGLNADEHMANTDSICQ PLIKGRSAPGRLNQPSAVGTRQ QKLTSAFQQQHLNFSQVQHFGS GHQEWNGNFGHRRQQAYIPTSV TSNPFTLSHGSPNHTAVHAHLA GNTHLGGQPTLLPYPSSAT

References:

Adra CN, Zhu S, Ko JL, Guillemot JC, Cuervo AM, Kobayashi H, Horiuchi T, Lelias JM, Rowley JD, Lim B.

LAPTM5: a novel lysosomal-associated multispanning membrane protein preferentially expressed in hematopoietic cells.
Genomics. 1996 Jul 15;35(2):328-37.

Allen M, Friedler A, Schon O, Bycroft M.

The Structure of an FF Domain from Human HYPA/FBP11.
J Mol Biol. 2002 Oct 25;323(3):411.

Allen RW, Trach KA, Hoch JA.

Identification of the 37-kDa protein displaying a variable interaction with the erythroid cell membrane as glyceraldehyde-3-phosphate dehydrogenase.
J Biol Chem. 1987 Jan 15;262(2):649-53.

Azuma T, Takei M, Yoshikawa T, Nagasugi Y, Kato M, Otsuka M, Shiraiwa H, Sugano S, Mitamura K, Sawada S, Masuho Y,Seki N.

Identification of candidate genes for Sjogren's syndrome using MRL/lpr mouse model of Sjogren's syndrome and cDNA microarray analysis.
Immunol Lett. 2002 May 1;81(3):171-6.

Bennett D, Alphey L.

PP1 binds Sara and negatively regulates Dpp signaling in Drosophila melanogaster.
Nat Genet. 2002 Aug;31(4):419-23.

Cabrita MA, Hobman TC, Hogue DL, King KM, Cass CE.

Mouse transporter protein, a membrane protein that regulates cellular multidrug resistance, is localized to lysosomes.
Cancer Res. 1999 Oct 1;59(19):4890-7.

Chalaux E, Lopez-Rovira T, Rosa JL, Bartrons R, Ventura F.

JunB is involved in the inhibition of myogenic differentiation by bone morphogenetic protein-2.

J Biol Chem. 1998 Jan 2;273(1):537-43.

Cohen PT.

Protein phosphatase 1--targeted in many directions.

J Cell Sci. 2002 Jan 15;115(Pt 2):241-56.

Dennler S, Itoh S, Vivien D, ten Dijke P, Huet S, Gauthier JM

Direct binding of Smad3 and Smad4 to critical TGF -inducible elements in the promoter of human palsminogen activator inhibitor type 1-gene.

EMBO J. 1998, 17:3091-3100

Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T.

Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells.

Nature. 2001 May 24;411(6836):494-8.

Faber PW, Barnes GT, Srinidhi J, Chen J, Gusella JF, MacDonald ME.

Huntingtin interacts with a family of WW domain proteins.

Hum Mol Genet. 1998 Sep;7(9):1463-74.

Goumans MJ, Valdimarsdottir G, Itoh S, Rosendahl A, Sideras P, ten Dijke P.

Balancing the activation state of the endothelium via two distinct TGF-beta type I receptors.

EMBO J. 2002 Apr 2;21(7):1743-53.

Hogue DL, Kerby L, Ling V.

A mammalian lysosomal membrane protein confers multidrug resistance upon expression in Saccharomyces cerevisiae.

J Biol Chem. 1999 Apr 30;274(18):12877-82.

Jiao K, Zhou Y, Hogan BL.

Identification of mZnf8, a Mouse Kruppel-Like Transcriptional Repressor, as a Novel Nuclear Interaction Partner of Smad1.

Mol Cell Biol. 2002 Nov;22(21):7633-44.

Kavsak P, Rasmussen RK, Causing CG, Bonni S, Zhu H, Thomsen GH, Wrana JL.

Smad7 binds to Smurf2 to form an E3 ubiquitin ligase that targets the TGF beta receptor for degradation.
Mol Cell. 2000 Dec;6(6):1365-75.

Keeton MR, Curriden SA, van Zonneveld AJ, Loskutoff DJ.

Identification of regulatory sequences in the type 1 plasminogen activator inhibitor gene responsive to transforming growth factor beta.
J Biol Chem. 1991 Dec 5;266(34):23048-52.

Kenny DA, Jurata LW, Saga Y, Gill GN.

Identification and characterization of LMO4, an LMO gene with a novel pattern of expression during embryogenesis.
Proc Natl Acad Sci U S A. 1998 Sep 15;95(19):11257-62.

Kost T, Condreay JP
Recombinant baculovirus as mammalian cell gene-delivery vectors
Trends in Biotechnology 2002, 20:173-180

Kusanagi K, Inoue H, Ishidou Y, Mishima K, Kawabata M, Miazono K
Characterization of a Bone Morphogenetic Protein-responsive Smad binding element.
Molecular Biology of the Cell 2000, 11:555-565

Lagali PS, Kakuk LE, Griesinger IB, Wong PW, Ayyagari R.

Identification and characterization of C6orf37, a novel candidate human retinal disease gene on chromosome 6q14.
Biochem Biophys Res Commun. 2002 Apr 26;293(1):356-65.

Lania L, Donti E, Pannuti A, Pascucci A, Pengue G, Feliciello I, La Mantia G, Lanfrancone L, Pelicci PG.

cDNA isolation, expression analysis, and chromosomal localization of two human zinc finger genes.
Genomics. 1990 Feb;6(2):333-40.

Mauviel A, Chung KY, Agarwal A, Tamai K, Uitto J.

Cell-specific induction of distinct oncogenes of the Jun family is responsible for differential regulation of collagenase gene expression by transforming growth factor-beta in fibroblasts and keratinocytes.
J Biol Chem. 1996 May 3;271(18):10917-23.

Mousses S, Bubendorf L, Wagner U, Hostetter G, Kononen J, Cornelison R, Goldberger N, Elkahloun AG, Willi N, Koivisto P, Ferhle W, Raffeld M, Sauter G, Kallioniemi OP.

Clinical validation of candidate genes associated with prostate cancer progression in the CWR22 model system using tissue microarrays.
Cancer Res. 2002 Mar 1;62(5):1256-60.

Origasa M, Tanaka S, Suzuki K, Tone S, Lim B, Koike T.

Activation of a novel microglial gene encoding a lysosomal membrane protein in response to neuronal apoptosis.
Brain Res Mol Brain Res. 2001 Mar 31;88(1-2):1-13.

Oshima A, Kyle JW, Miller RD, Hoffmann JW, Powell PP, Grubb JH, Sly WS, Tropak M, Guise KS, Gravel RA.

Cloning, sequencing, and expression of cDNA for human beta-glucuronidase.
Proc Natl Acad Sci U S A. 1987 Feb;84(3):685-9.

Passani LA, Bedford MT, Faber PW, McGinnis KM, Sharp AH, Gusella JF, Vonsattel JP, MacDonald ME.

Huntingtin's WW domain partners in Huntington's disease post-mortem brain fulfill genetic criteria for direct involvement in Huntington's disease pathogenesis.
Hum Mol Genet. 2000 Sep 1;9(14):2175-82.

Patel PI, Framson PE, Caskey CT, Chinault AC.

Fine structure of the human hypoxanthine phosphoribosyltransferase gene.
Mol Cell Biol. 1986 Feb;6(2):393-403.

Seki N, Hattori A, Hayashi A, Kozuma S, Sasaki M, Suzuki Y, Sugano S, Muramatsu MA, Saito T.

Cloning and expression profile of mouse and human genes, Rnf11/RNF11, encoding a novel RING-H2 finger protein.

Biochim Biophys Acta. 1999 Dec 23;1489(2-3):421-7.

Schmittgen TD, Zakrajsek BA.

Effect of experimental treatment on housekeeping gene expression: validation by real-time, quantitative RT-PCR.

J Biochem Biophys Methods. 2000 Nov 20;46(1-2):69-81.

Scott LM, Mueller L, Collins SJ.

E3, a hematopoietic-specific transcript directly regulated by the retinoic acid receptor alpha.

Blood. 1996 Oct 1;88(7):2517-30.

Sum EY, Peng B, Yu X, Chen J, Byrne J, Lindeman GJ, Visvader JE.

The LIM domain protein LMO4 interacts with the cofactor CtIP and the tumor suppressor BRCA1 and inhibits BRCA1 activity.

J Biol Chem. 2002 Mar 8;277(10):7849-56.

ten Dijke P, Yamashita H, Ichijo H, Franzen P, Laiho M, Miyazono K, Heldin CH.

Characterization of type I receptors for transforming growth factor-beta and activin.
Science. 1994 Apr 1;264(5155):101-4.

Tse E, Grutz G, Garner AA, Ramsey Y, Carter NP, Copeland N, Gilbert DJ, Jenkins NA, Agulnick A, Forster A, Rabbitts TH.

Characterization of the Lmo4 gene encoding a LIM-only protein: genomic organization and comparative chromosomal mapping.
Mamm Genome. 1999 Nov;10(11):1089-94.

Visvader JE, Venter D, Hahm K, Santamaria M, Sum EY, O'Reilly L, White D, Williams R, Armes J, Lindeman GJ.

The LIM domain gene LMO4 inhibits differentiation of mammary epithelial cells in vitro and is overexpressed in breast cancer.
Proc Natl Acad Sci U S A. 2001 Dec 4;98(25):14452-7.

Wagner EF, Karsenty G.

Genetic control of skeletal development.
Curr Opin Genet Dev. 2001 Oct;11(5):527-32.

Walker MG.

Z39Ig is co-expressed with activated macrophage genes.
Biochim Biophys Acta. 2002 Apr 12;1574(3):387-90.

Watanabe T, Ichihara M, Hashimoto M, Shimono K, Shimoyama Y, Nagasaka T, Murakumo Y, Murakami H, Sugiura H, Iwata H, Ishiguro N, Takahashi M.

Characterization of gene expression induced by RET with MEN2A or MEN2B mutation.
Am J Pathol. 2002 Jul;161(1):249-56.

Zawel L, Dai JL, Buckhaults P, Zhou S, Kinzler KW, Vogelstein B, Kern SE
Human Smad3 and Smad4 are sequence specific transcription activators
Molecular Cell 1998, 1:611-617

CLAIMS

What is claimed is:

1. A complex between two interacting proteins as defined in columns 1 and 4 in Table 2.
2. A complex between two polynucleotides encoding for the polypeptides of claim 1.
3. A recombinant host cell expressing the interacting polypeptides of said complex of protein-protein interaction of claim 1.
4. Use of a SID®, an interaction or a prey to screen molecules that inhibit TGFβ or inhibit a TGFβ super-family of cytokines pathway.
5. A molecule that inhibits inhibit TGFβ or inhibits a TGFβ super-family of cytokines pathway.
6. Use according to Claim 4, wherein said screening occurs in mammalian cells or yeast cells.
7. A SID® polypeptide comprising the SEQ ID No 63 to 98.
8. A SID® polynucleotide comprising the SEQ ID No 27 to 62.
9. A vector comprising the SID® polynucleotide comprising the SEQ ID No 27 to 62.
10. A fragment of said SID® polypeptide according to Claim 7.
11. A variant of said SID® polypeptide according to Claim 7.
12. A fragment of said SID® polynucleotide according to Claim 8.
13. A variant of said SID® polynucleotide according to Claim 8.
14. A vector comprising the SID® polynucleotide according to any one of Claims 8, 12 or 13.
15. A recombinant host cell containing the vectors according to Claim 14.
16. A pharmaceutical composition comprising the molecule of claim 5 and a pharmaceutically acceptable carrier.
17. A pharmaceutical composition comprising a SID® polypeptide SEQ ID No 63 to 98 and a pharmaceutically acceptable carrier.
18. A pharmaceutical composition comprising the recombinant host cells of Claim 15 and a pharmaceutically acceptable carrier.
19. A protein chip comprising the polypeptides of Table 2.
20. Use of a ZNF8 protein for the preparation of a medicament for treating diseases and /or disorders linked or involving a TGFβ super-family of cytokines.

21. Use of a LAPTm5 protein for the preparation of a medicament for treating diseases and /or disorders linked or involving a TGFβ super-family of cytokines.

22. Use of a RNF11 protein for the preparation of a medicament for treating diseases and /or disorders linked or involving a TGFβ super-family of cytokines.

23. Use of a LMO4 protein for the preparation of a medicament for treating prostate cancer.

24. Use of a PPC1 protein for the preparation of a medicament for treating diseases and /or disorders linked or involving a TGFβ super-family of cytokines.

25. Use of an HYPA protein for the preparation of a medicament for treating diseases and /or disorders linked or involving a TGFβ super-family of cytokines.

26. Use of a PTP protein for the preparation of a medicament for treating diseases and /or disorders linked or involving a TGFβ super-family of cytokines.

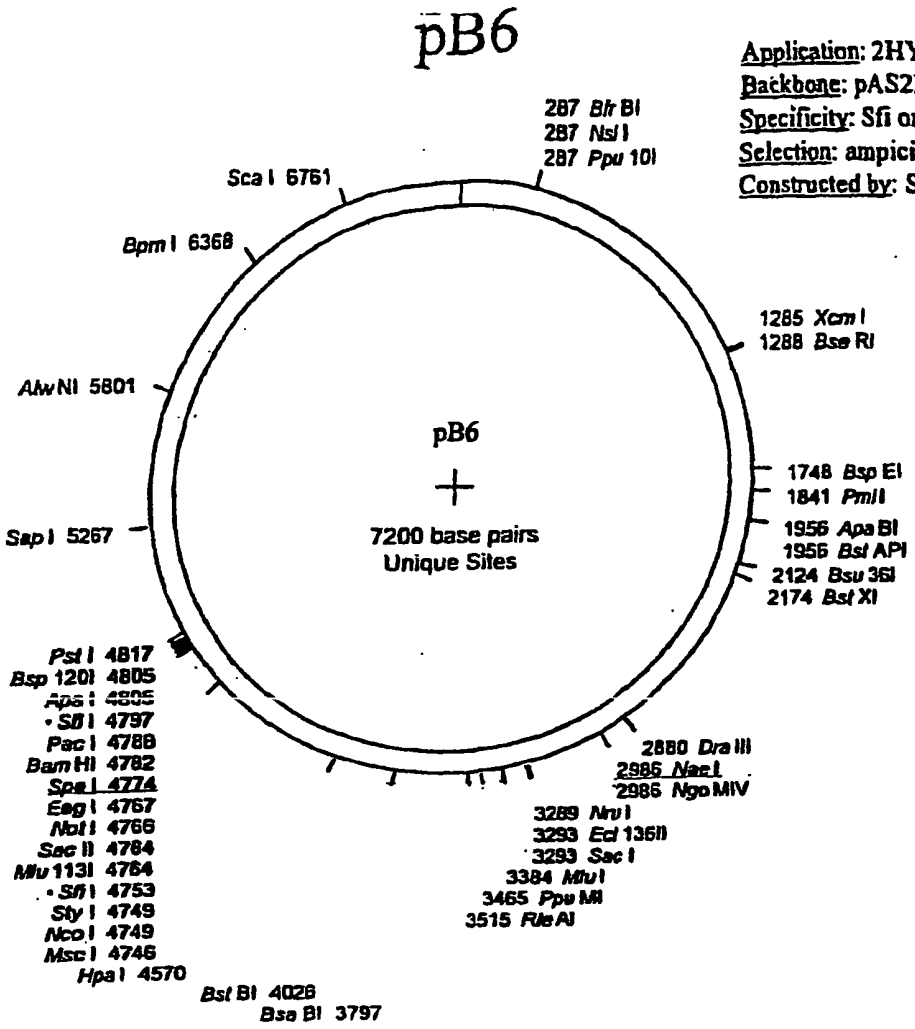
27. Use of an HYPK3 protein for the preparation of a medicament for treating diseases and /or disorders linked or involving a TGFβ super-family of cytokines.

28. Use of a KIAA1196 protein for the preparation of a medicament for treating diseases and /or disorders linked or involving a TGFβ super-family of cytokines.

29. Use of a FL20037 protein for the preparation of a medicament for treating diseases and /or disorders linked or involving a TGFβ super-family of cytokines.

30. Use of a complex between two interacting proteins as defined in columns 1 and 4 in table 2 to screen I molecules for diagnosis or treating transforming growth factor β disorders and/or diseases.

FIG. 1



Oligo 160

gagagtagtaacaaaggtcAAAGACAGTTGACTGTATCGCCG GAA TTT ATG

Sfi I Sac II Spe I Bam HI

GCC ATG GCC GGA CGG GCC GCG GCC GCA CTA GTG GGG ATC C

Nco I Not I

STOP Sfi I Apa I Pst I

TT AAT TAA GGG CCA CTG GGG CCC CTC GAC CTG CAG CCA

Pac I

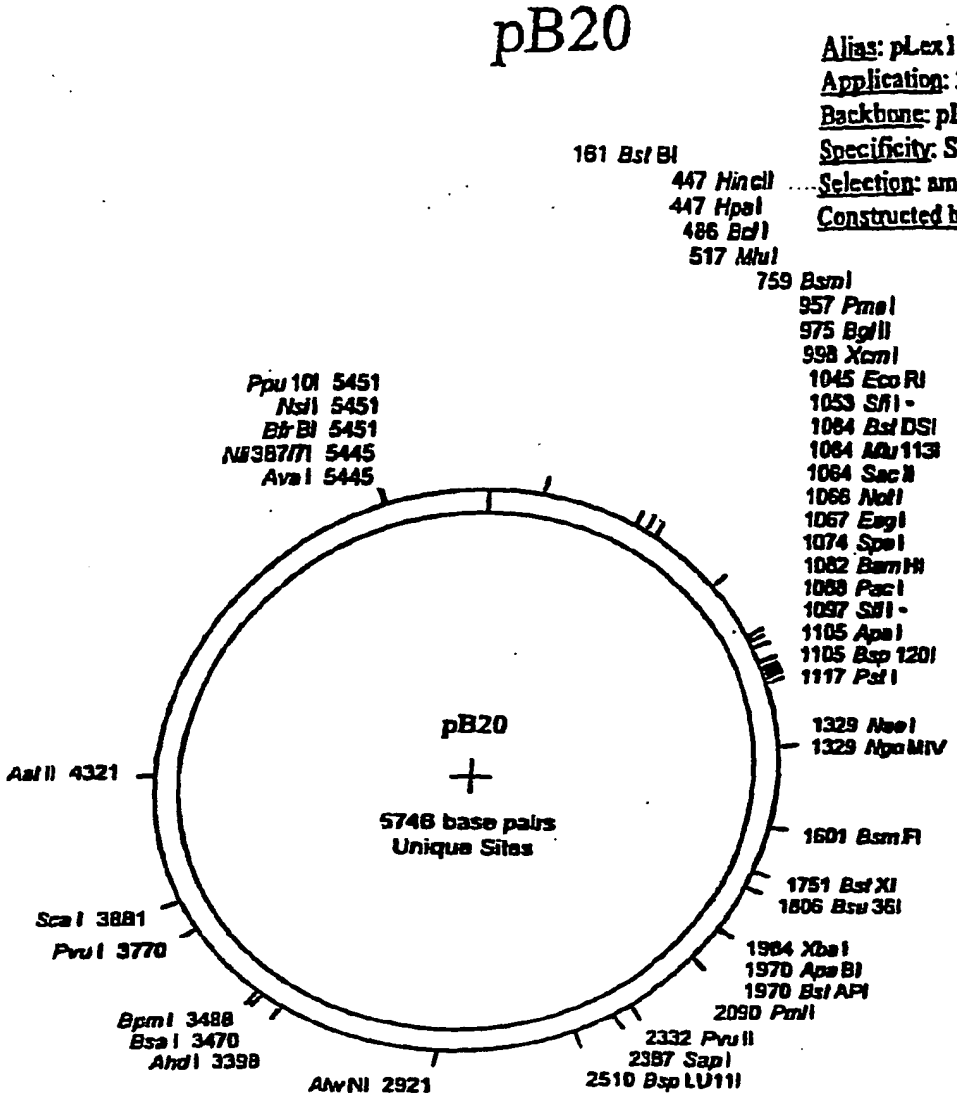
Oligo 161

AGC TAA TTccgggcgaatttcttatg

Oligo 160 5' GAGAGTAGTAACAAAGGTC3'

Oligo 161 5' CATAAGAAATTCGCCCGG 3'

FIG. 2



EcoR I — Sfi I — Not I — Spe I

GAA TTC GGG GCC GGA CGG GCC GCG GCC GCA CTA GTG

Sac II

BamH I — STOP

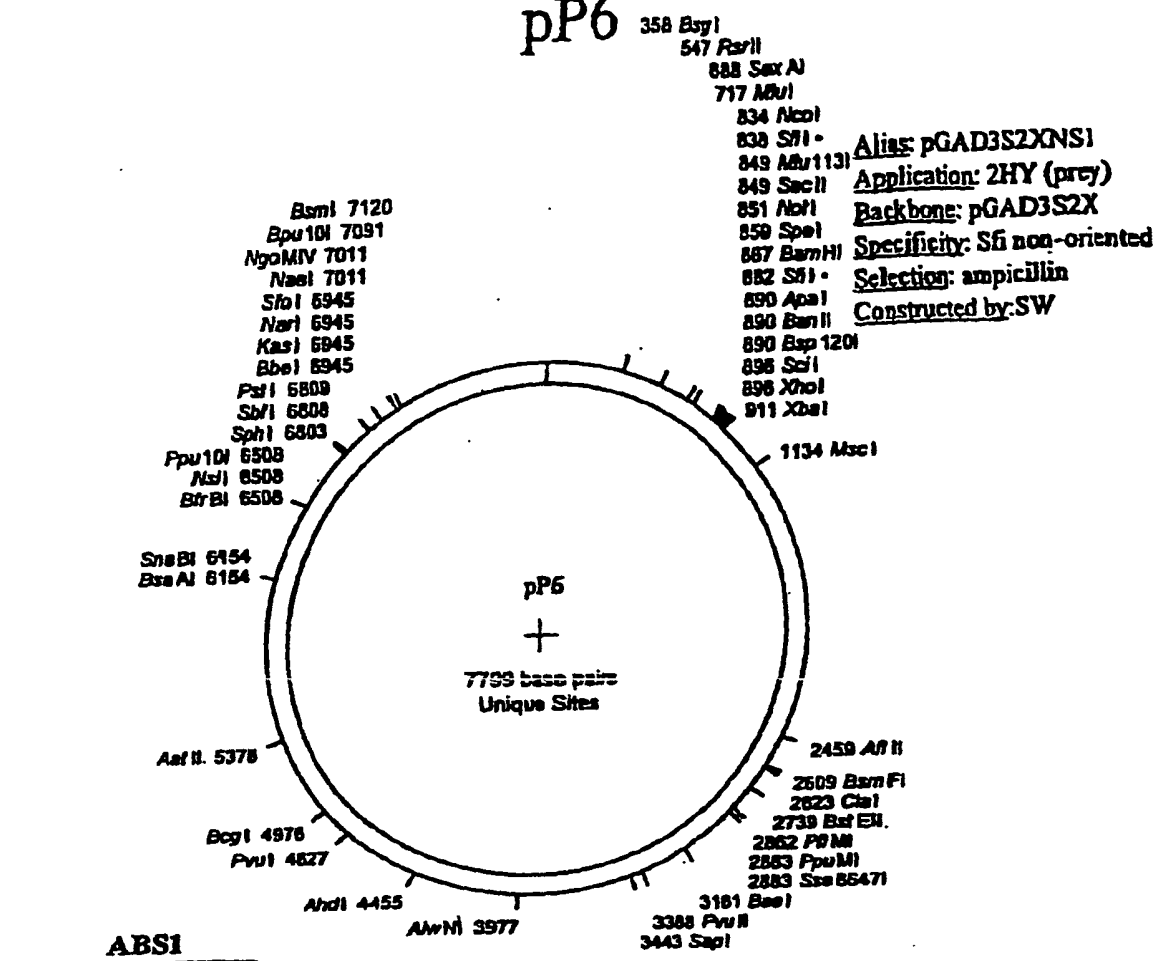
GGG ATC CTT AAT TAA GGG CCA CTG GGG CCC CTC GAC

Pac I — Sfi I

CTG CAG

Pst I

FIG. 3



BEST AVAILABLE COPY

ABS1

cgtttggaatcactacagg GATGTTTAATACCACTACAATGGATGATGTATATAACTATCTATT

JC90

cgatgatgaagataccccaccaaa CCCAAAAAAAAGAGATCCTAGAACTA

Sfi I **Sac II** **Spe I** **Bam HI**

GCC ATG GCC GCA GGG GCC GCG GCC GCA CTA GTG GGG ATC C

Nco I **Not I**

STOP **Sfi I** **Xho I** **Xba I**

TT AAT TAA GGG CCA CTG GGG CCC CTC GAG TAG CTA GTG TCT AGA

STOP **STOP** **STOP**

GGCCCGGTACCCAATTCGCCCTATAGTGAGTCGTATTACAATTCACTGGCCG TCGTTTTA

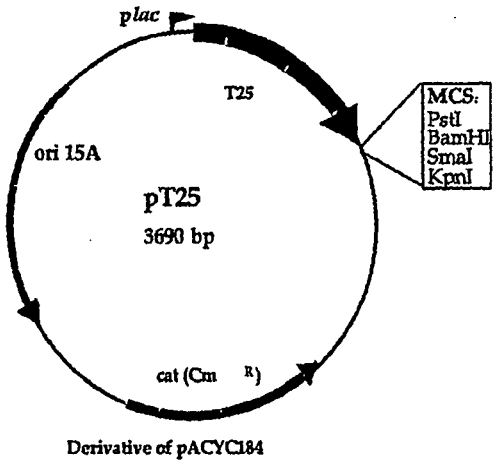
CAACGTCGTGACTGGGAAAACCCTGATCTATGAAT cgtagatactgaaaaacccc GCAA

162

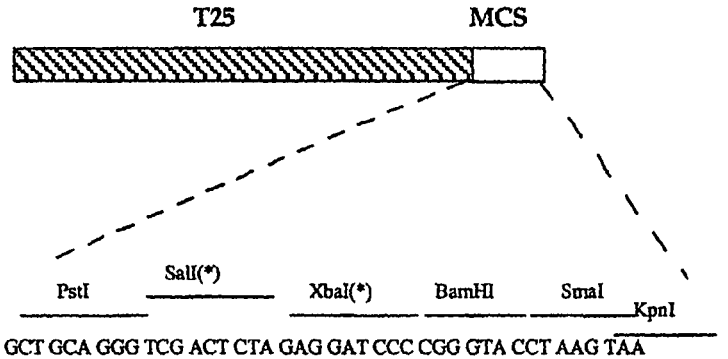
GTT cacttcaactgtgcatcgtg caccatctcaatttctttc

ABS2 **53**

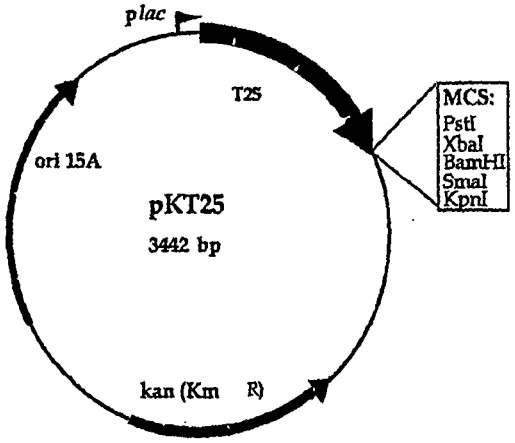
ABS1 5' CGTTTGGAATCACTACAGG 3'
JC90 5' CGATGATGAAGATACCCCACCAAA 3'
162 5' GGGGTTTTTCAGTATCTACG 3'
ABS2 5' CACGATGCACAGTTGAAGTG 3'
53 5' GAAATTGAGATGGTGCACGATGCAC 3'



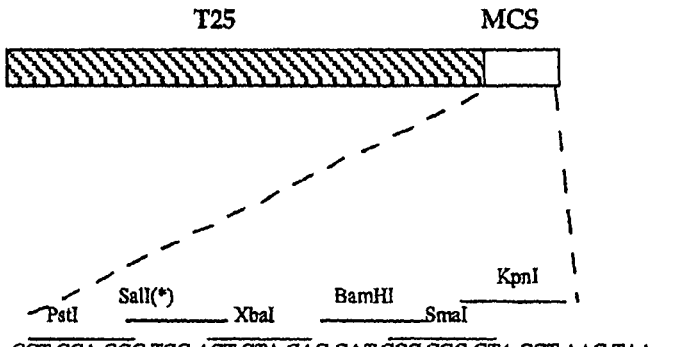
Derivative of pACYC184



(*) Restriction site is not unique

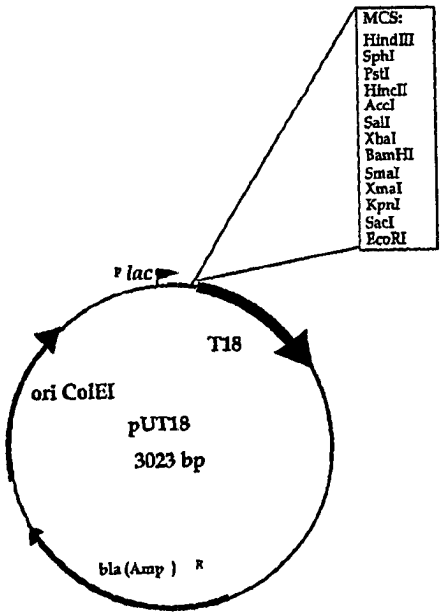


Derivative of pSU40

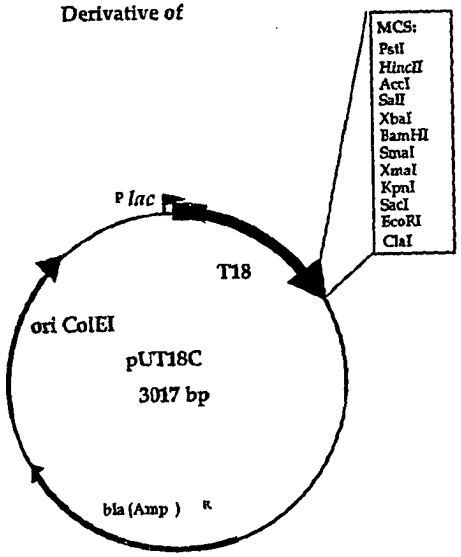


(*) Restriction site is not unique

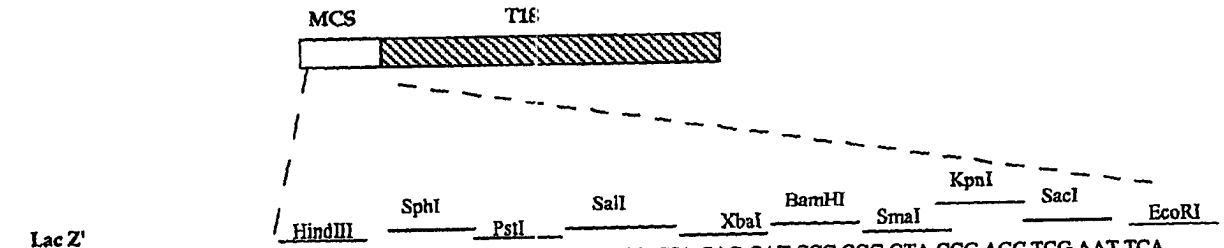
Figure 4 : Vectors expressing the T25 fragment



Derivative of



Derivative of pUC19



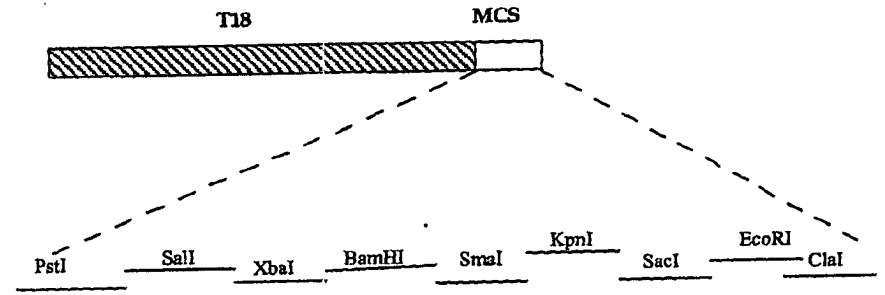


Figure 5

FIG. 6

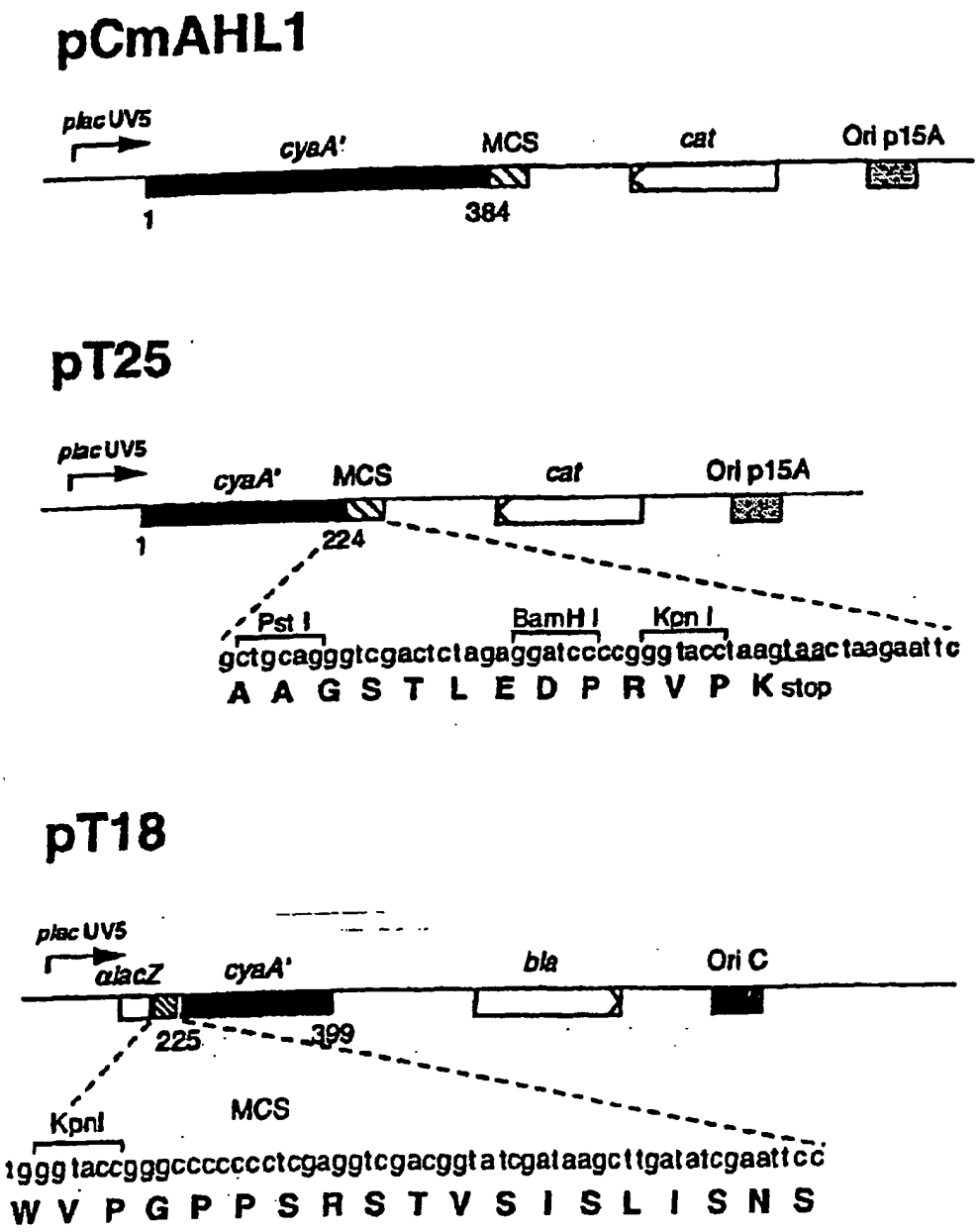


FIG. 7

Selectide Interaction Domain (SID®)

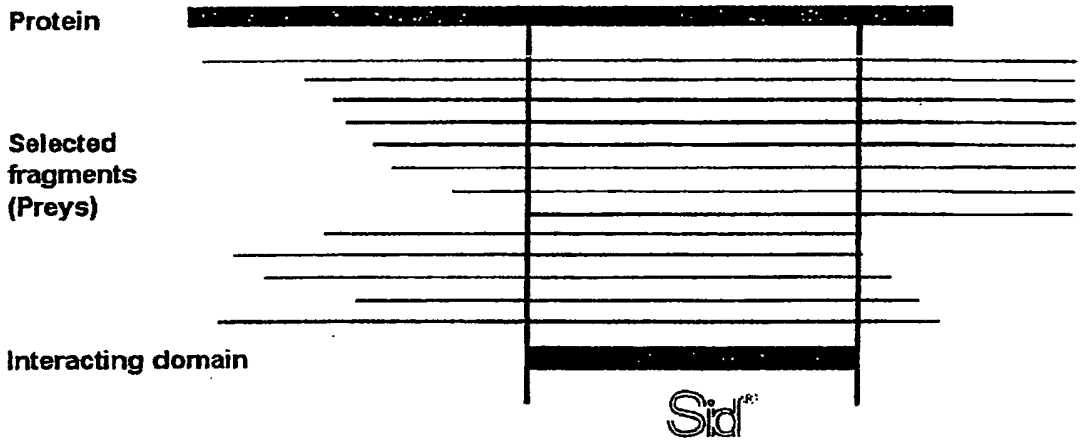
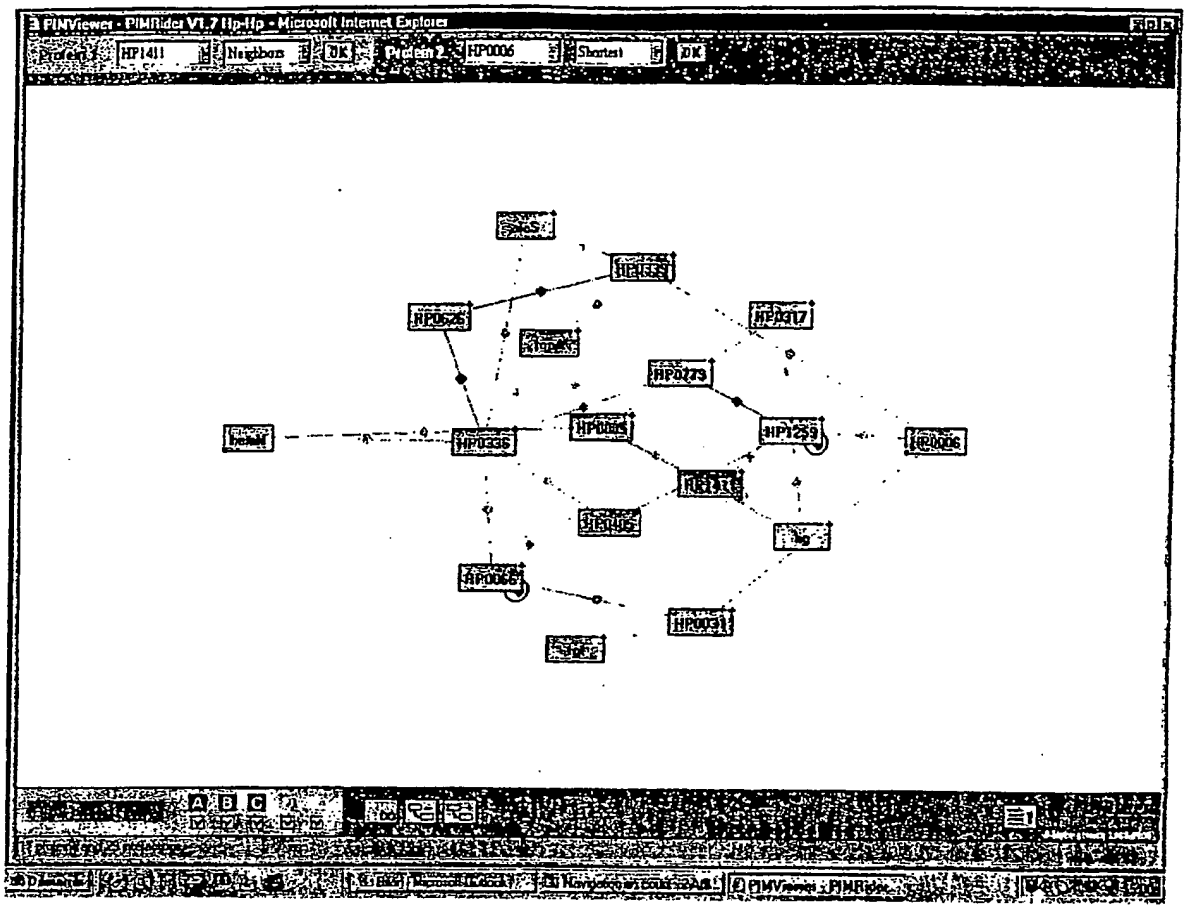
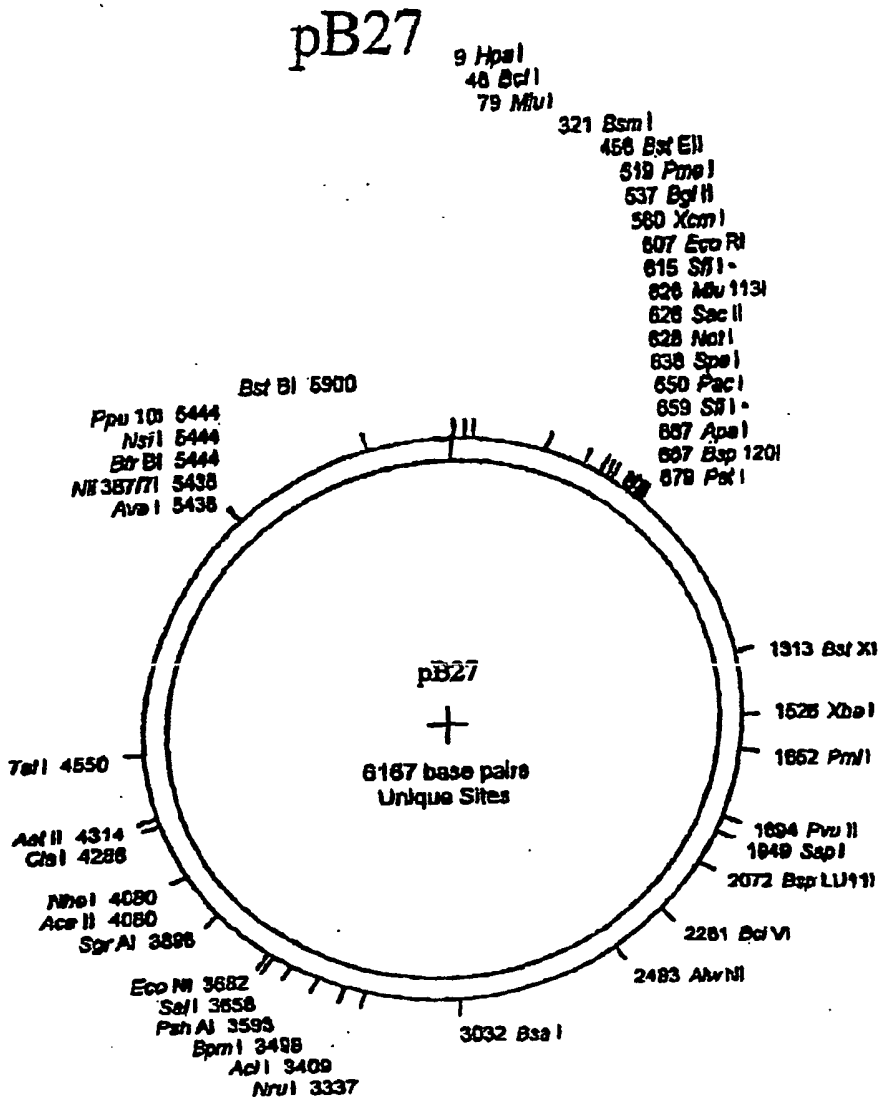


FIG. 8



Example of Protein Interaction Map

FIG. 9



oli946

aaagtcgaac tgttgccagaaaatagcgag tttaaaccaattgtcgtaatcttcgtca
gcagagcttcaccattgaagggctggcggttggggttattcgcaacggcgactggctg

EcoR I — Sfi I — Sac II — Spe I

GAA TTC GGG GCC GGA CGG GCC GCG GCC GCA CTA GTG GGG ATC

Not I

STOP — Sfi I — Apa I — Pst I

CTT AAT TAA GGG CCA CTG GGG CCC CTC GAC CTG CAG

Pac I

oli947

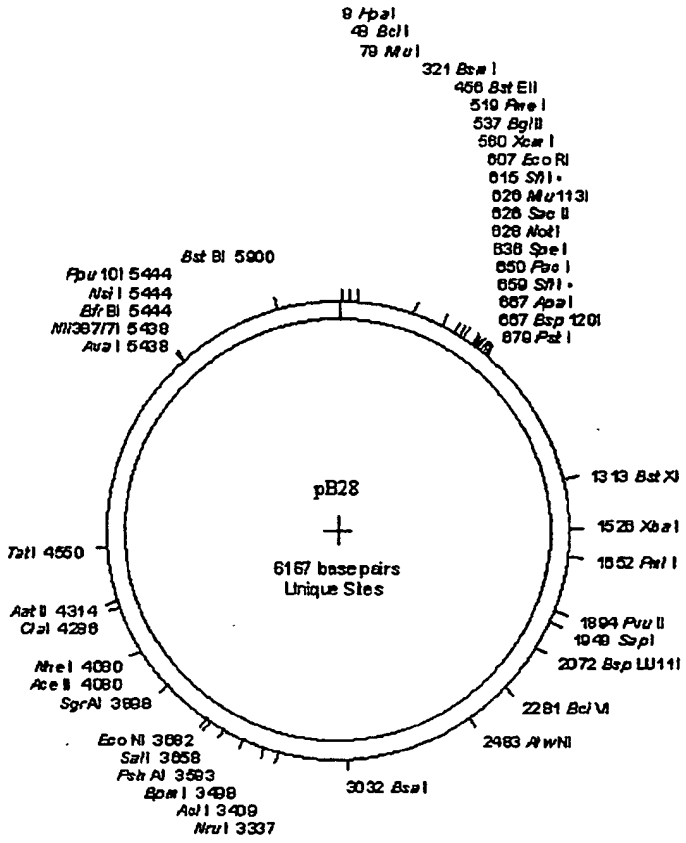
ccaa gctaattccgggcgaatt tcttatgatttatgatttttattattaaataagttataaaaaaaataa

Oligo 946 5' TGTTGCCAGAAAATAGCGAG 3'
Oligo 947 5' AATTCGCCCGGAATTAGC 3'

Figure 10

pB28

Alias: pLexTetNS1
Application: 2HY (bait)
Backbone: pB8
Specificity: Sfi non-oriented
Selection: tetracyclin
Constructed by: CR



EcoR I — Sfi I — Sac II — Spe I

GAA TTC GGG GCC GCA GGG GCC GCG GCC GCA CTA GTG GGG ATC

Not I

STOP — Sfi I — Apa I — Pst I

CTT AAT TAA GGG CCA CTG GGG CCC CTC GAC CTG CAG

Pac I

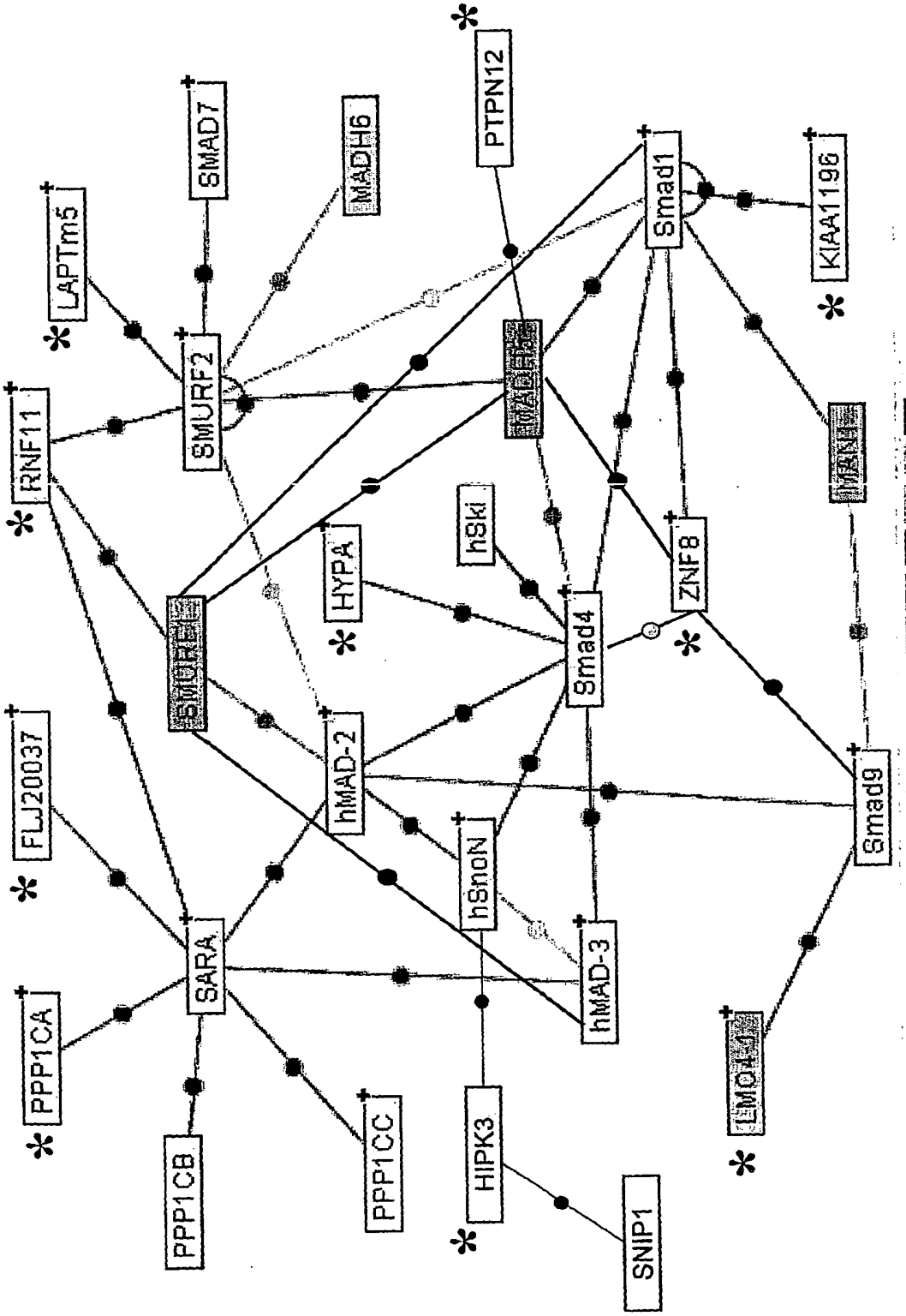


Figure 11: Protein interaction map around the newly functionally characterized proteins (*)

BEST AVAILABLE COPY

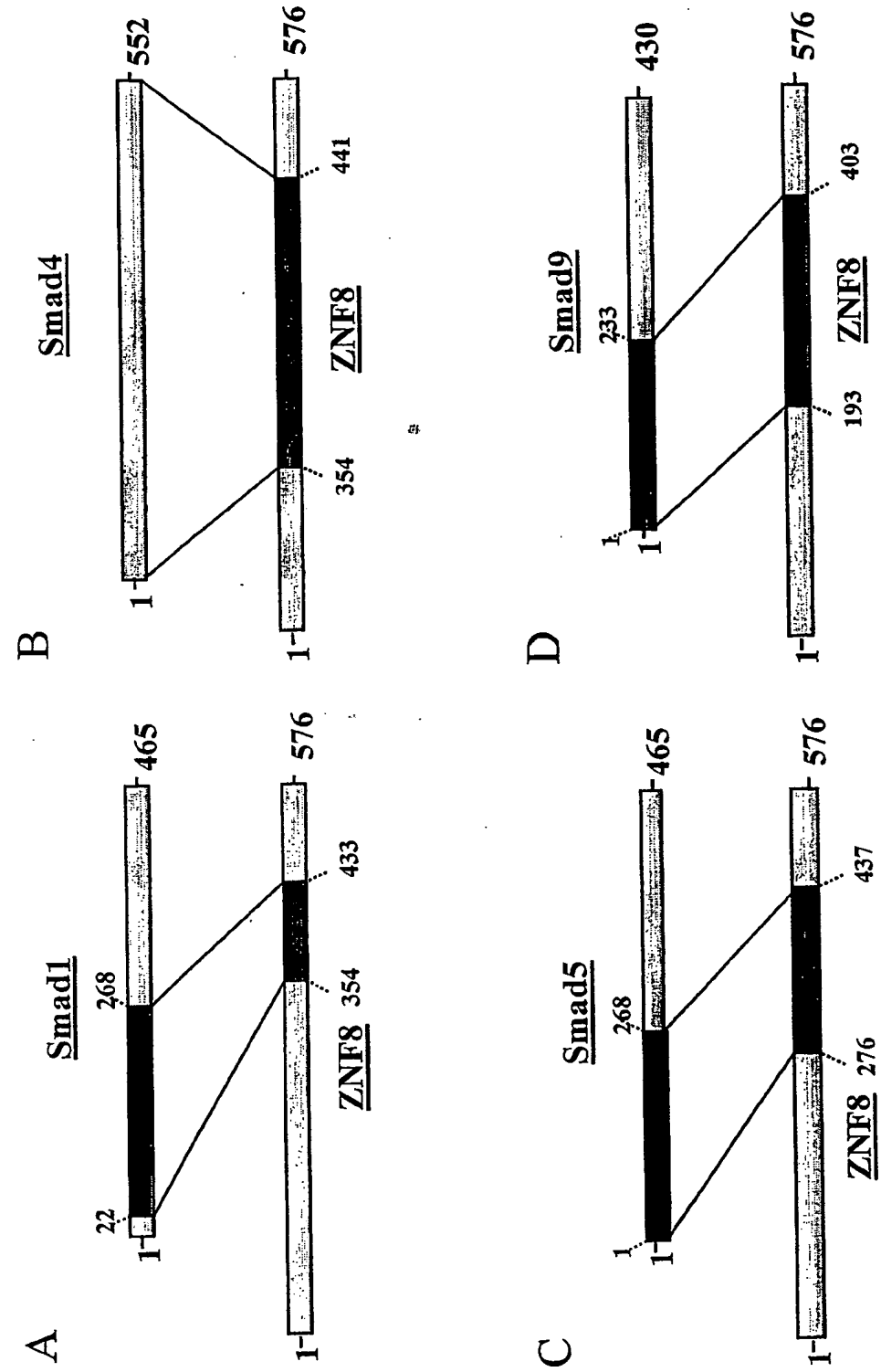
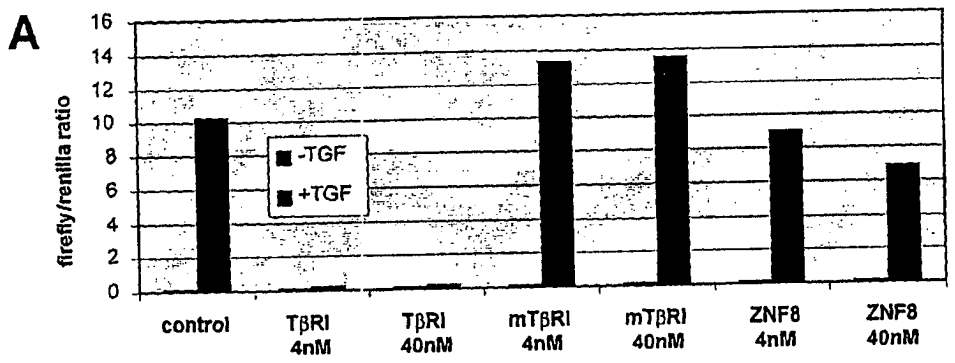


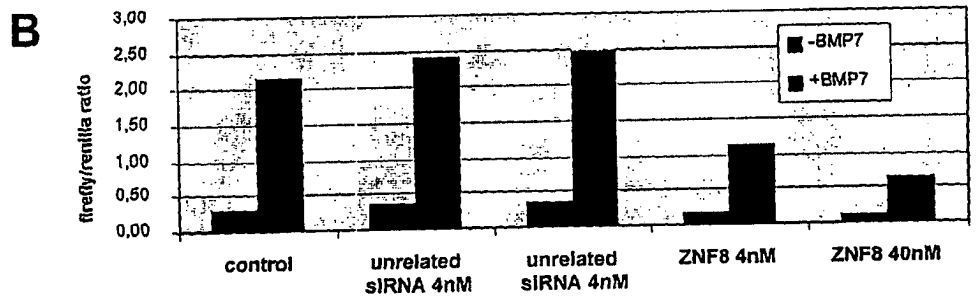
Figure 12: Protein Interaction map around ZNF8

BEST AVAILABLE COPY

ZNF8 siRNA regulation of $p(GTCT)_8$-MLP-Luc activity in HepG2 cells



ZNF8 siRNA regulation of $p(GC)_{12}$-MLP-Luc activity in HepG2 cells



ZNF8 siRNA regulation of $p(GC)_{12}$-MLP-Luc activity in HepG2 cells

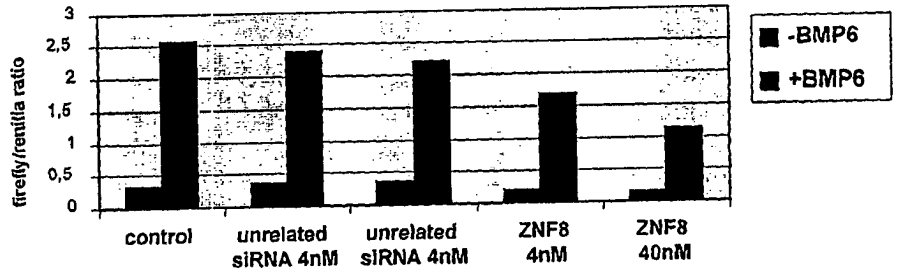


Figure 13: ZNF8 siRNA represses TGFβ- and BMP-dependent luciferase reporter activities.

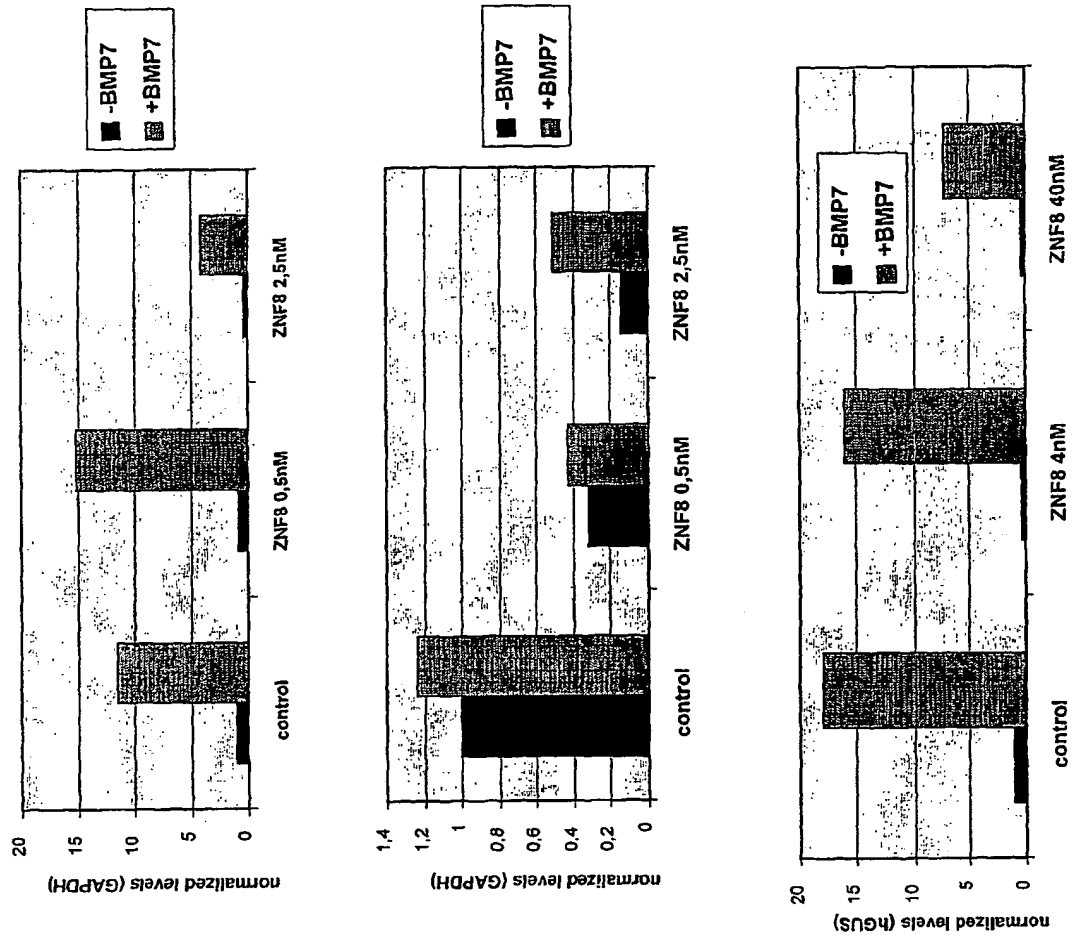
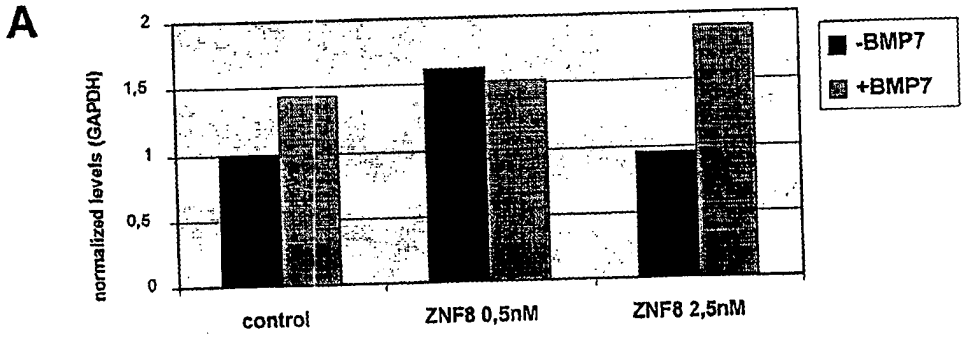


Figure 14: ZNF8 siRNA specifically represses BMP-dependent markers

Q-PCR analysis of PAI-1 mRNA levels in HepG2 cells



Q-PCR analysis of hGUS mRNA levels in HepG2 cells

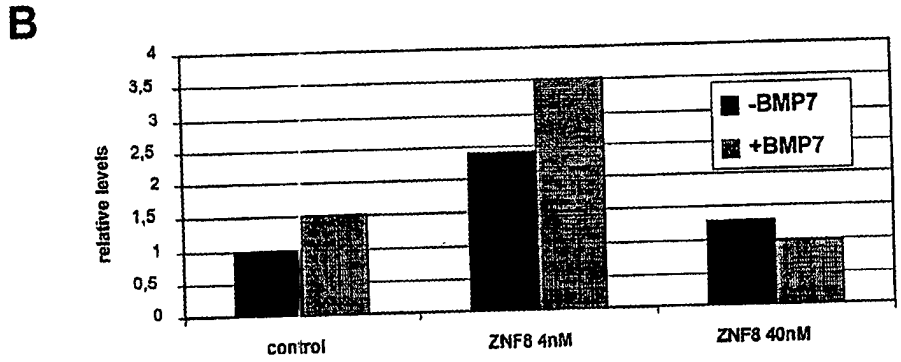


Figure 15: ZNF8 siRNA does not repress BMP-independent markers

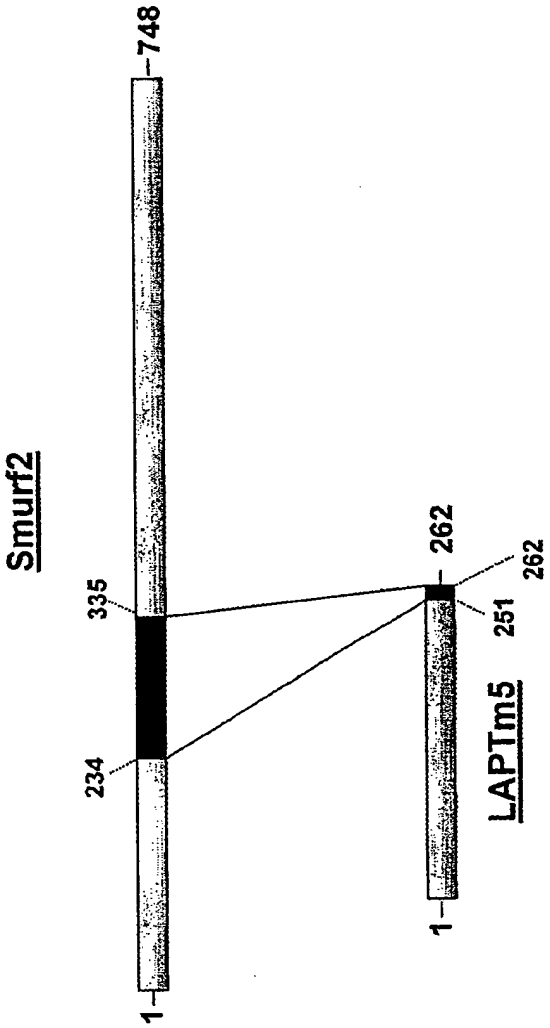


Figure 16: Interaction between LAPTm5 and Smurf2

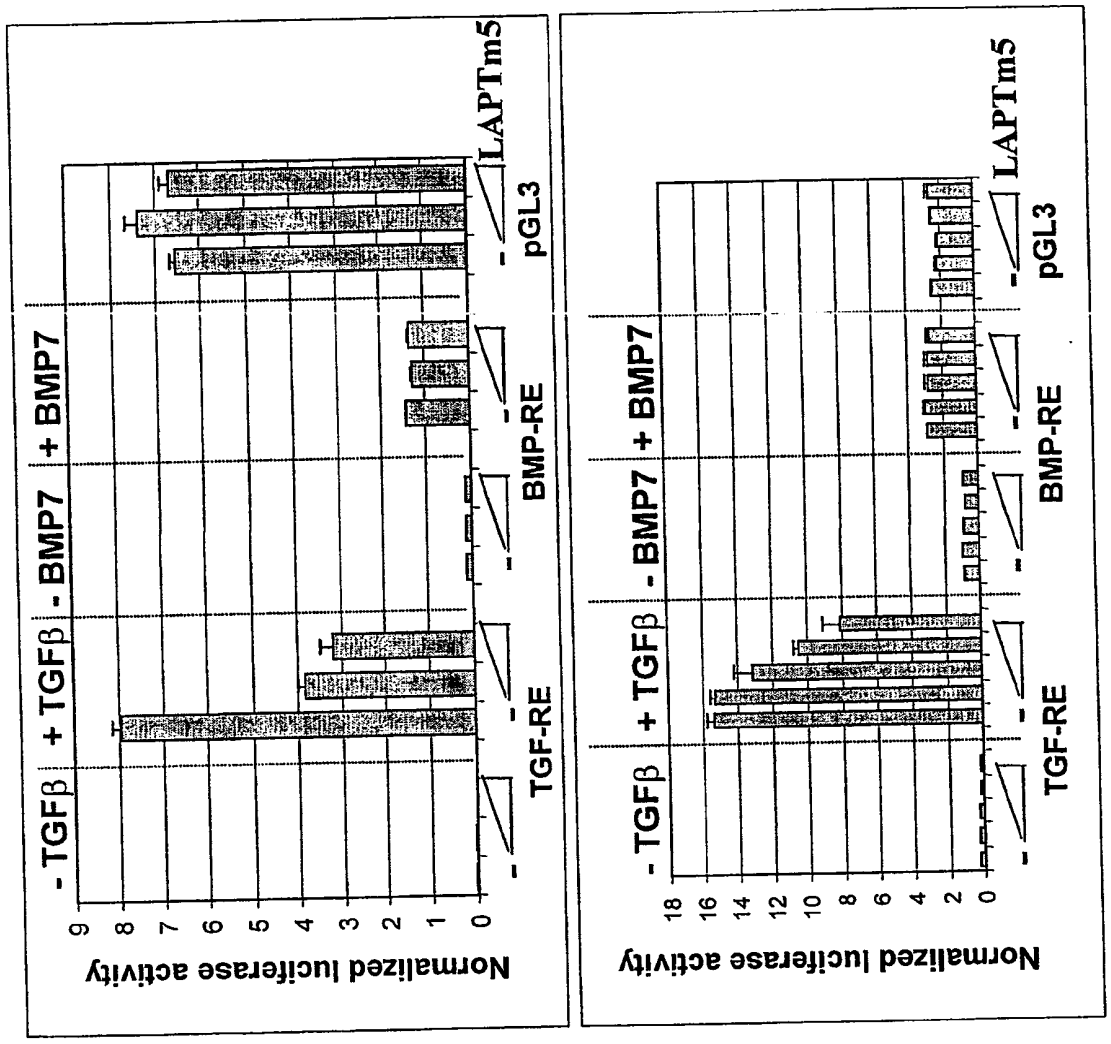


Figure 17: LAPTm5 specifically inhibits the TGFβ pathway

A

Cell lines	Ct level (LAPTm5)
HepG2	38
Hela	40
WI38	37,2
CEM	24,9
CEMC7	25,4
K562	30,7
JURKAT	25,2

B

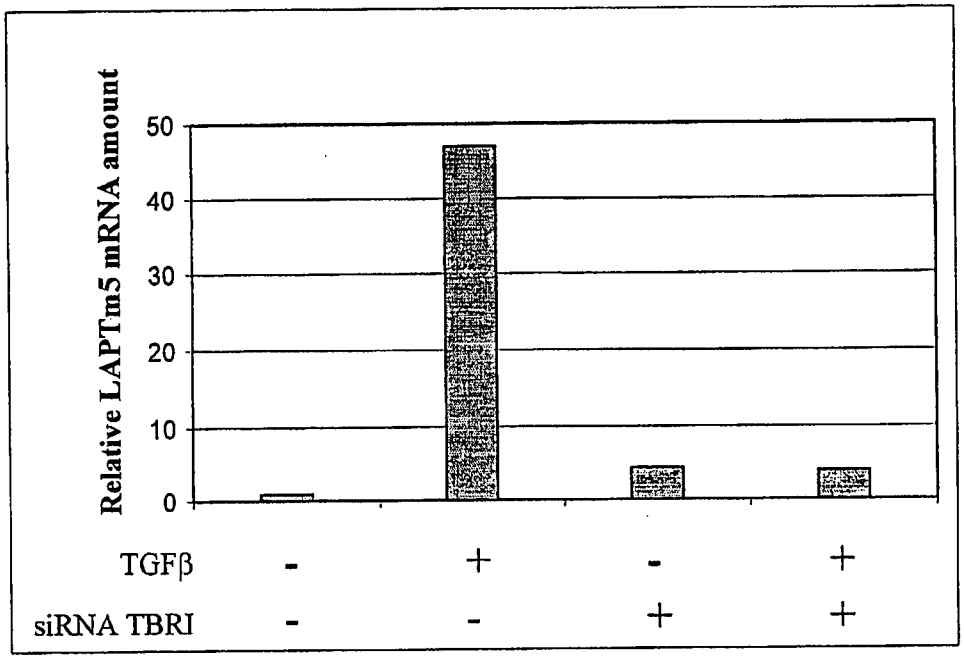


Figure 18: LAPTm5 expression is up-regulated by TGFβ

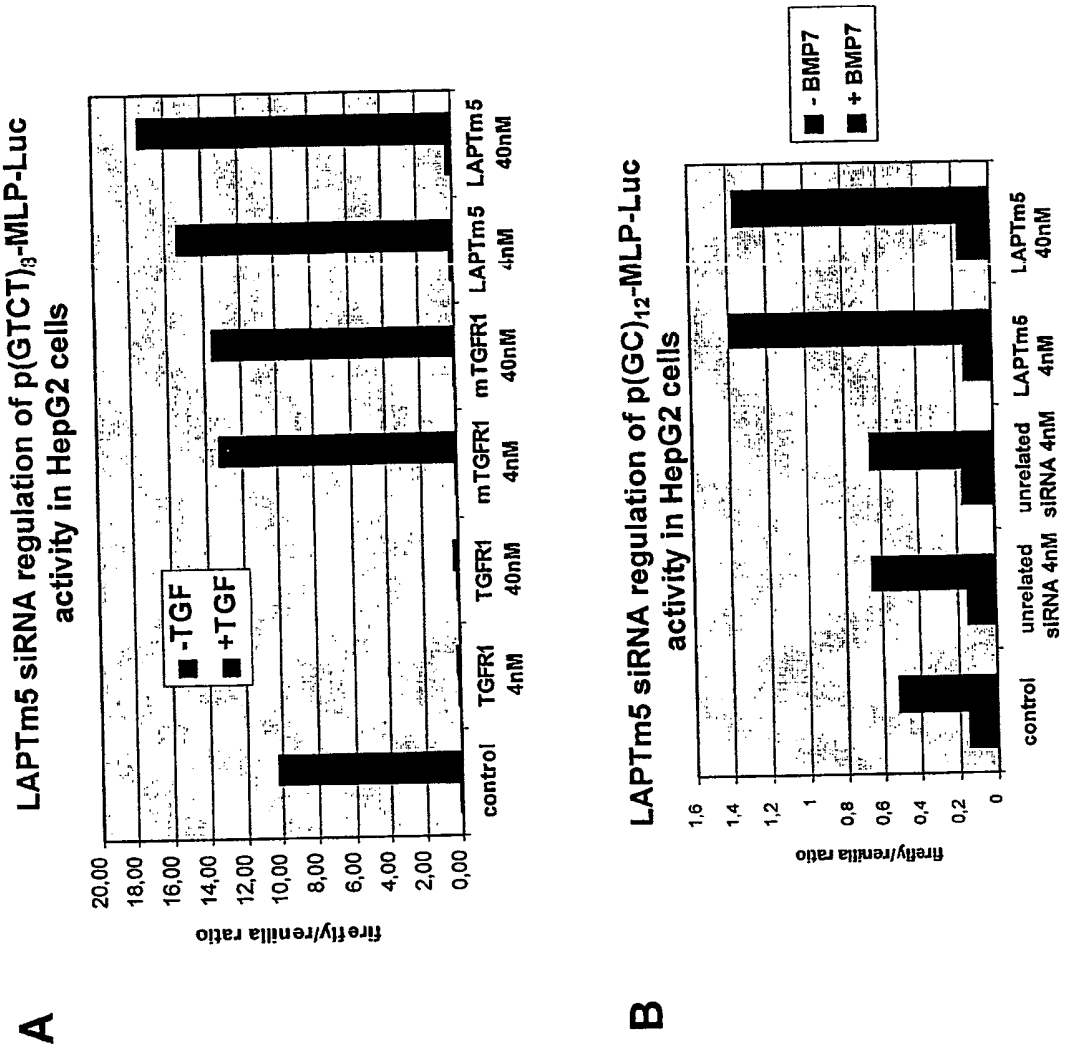


Figure 19: LAPTm5 siRNA up-regulates BMP and TGFβ-dependent reporter activities

BEST AVAILABLE COPY

Q-PCR analysis of PAI-1 mRNA levels in HepG2 cells

Q-PCR analysis of junB mRNA levels in HepG2 cells

Q-PCR analysis of alkaline phosphatase mRNA levels in HepG2 cells

Q-PCR analysis of hGUS mRNA levels in HepG2 cells

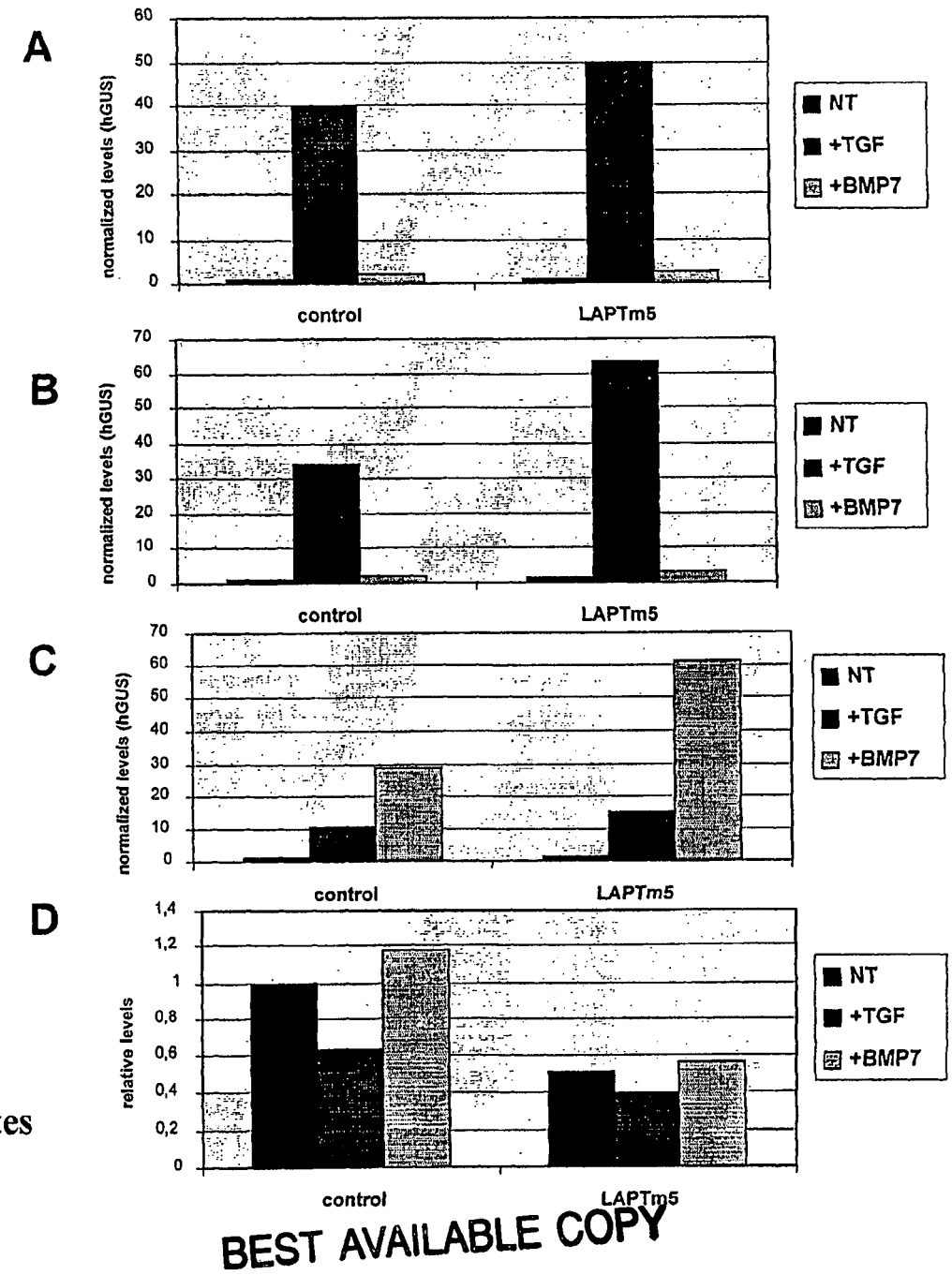


Figure 20: LAPTm5 siRNA up-regulates BMP and TGFβ-dependent markers

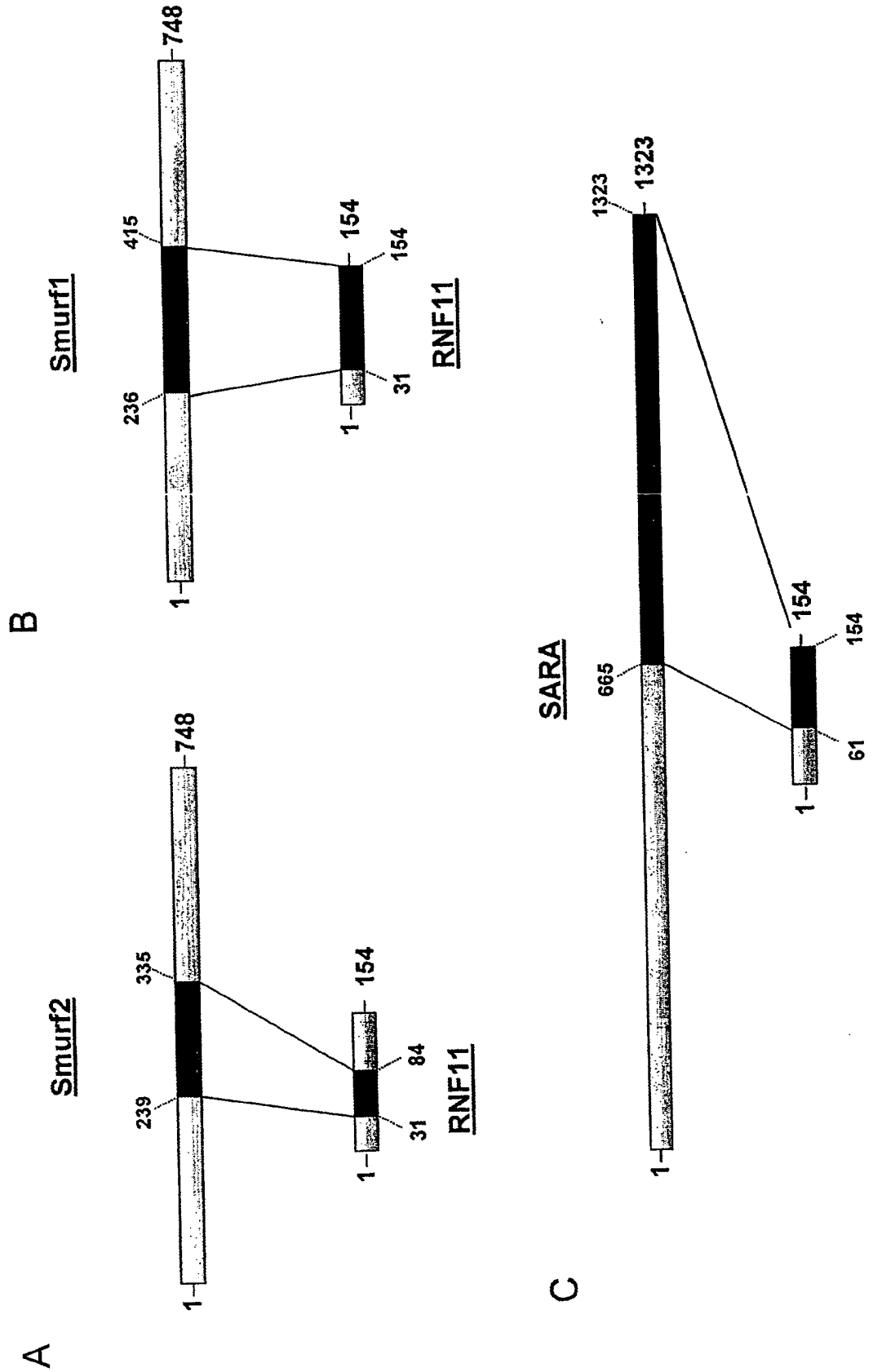


Figure 21: Interaction between RNF11 Smurf1, Smurf2 and SARA

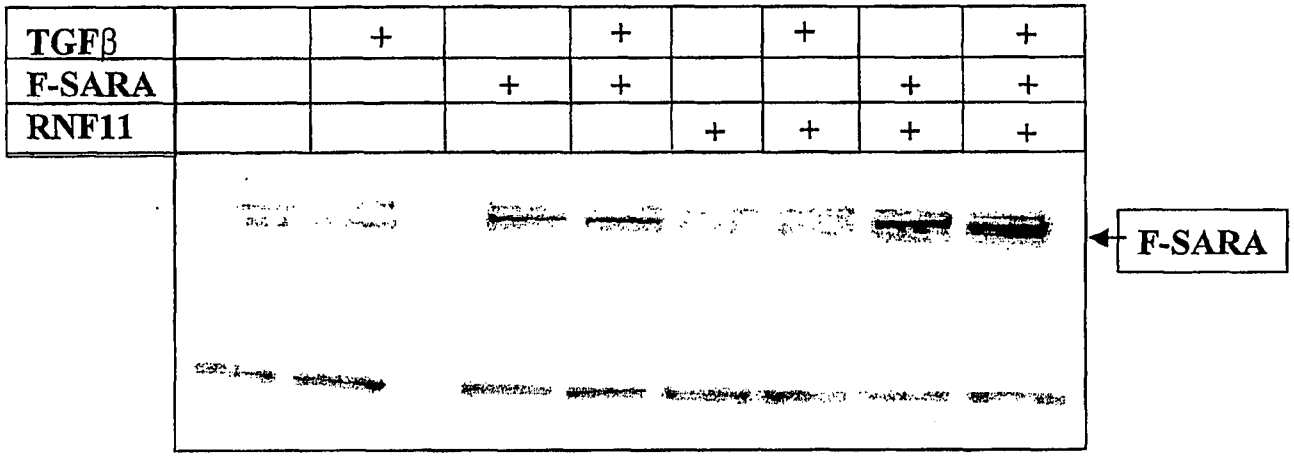


Figure 22: RNF11 is involved in regulating SARA protein levels

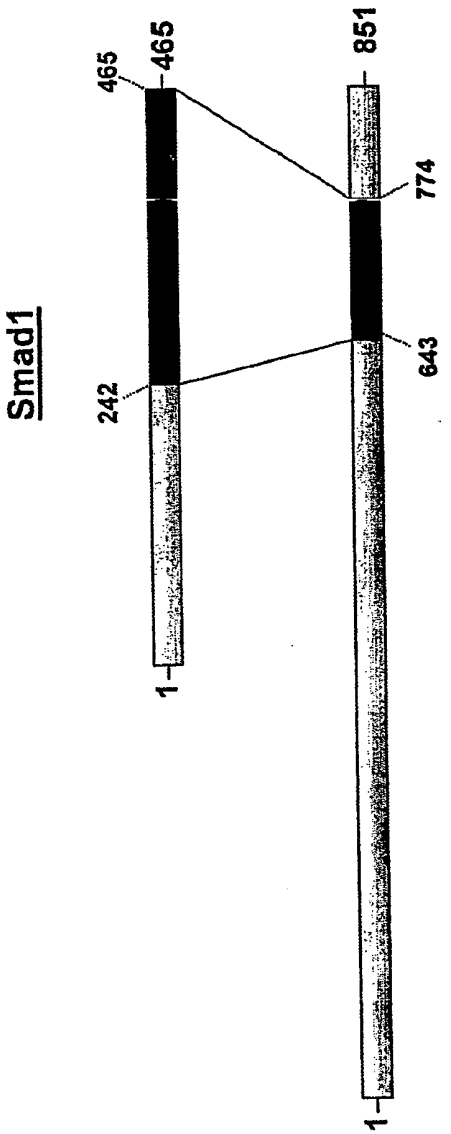
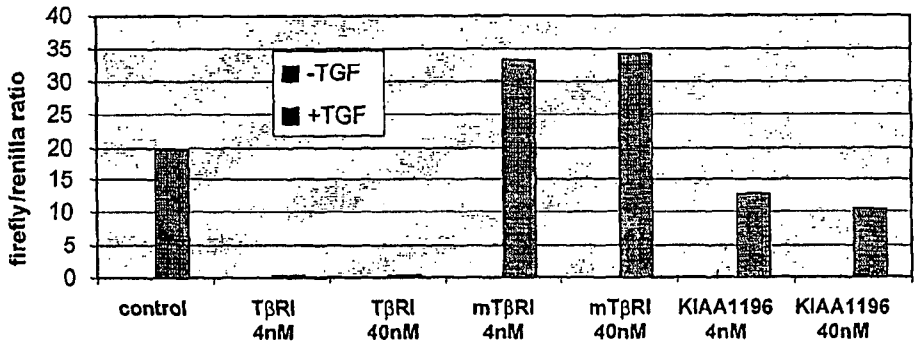


Figure 23: Interaction between KIAA1196 and Smad1

A

KIAA1196 siRNA regulation of $p(GTCT)_8$-MLP-Luc activity in HepG2 cells



B

KIAA1196 siRNA regulation of $p(GC)_{12}$-MLP-Luc activity in HepG2 cells

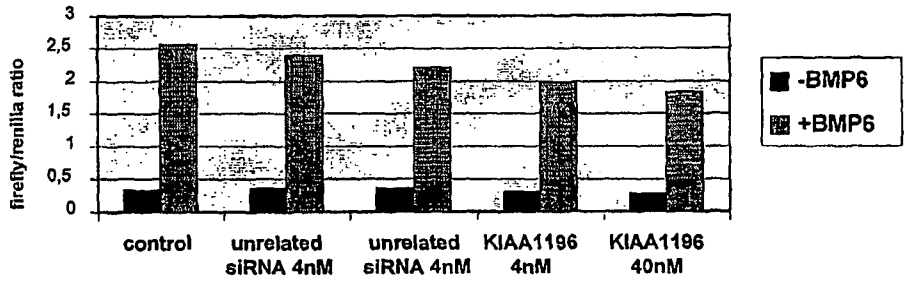


Figure 24: KIAA1196 siRNA specifically represses TGFβ-dependent reporter activity in HepG2 cells

KIAA1196 siRNA regulation of pGTCT-luc activity in HEK293 cells

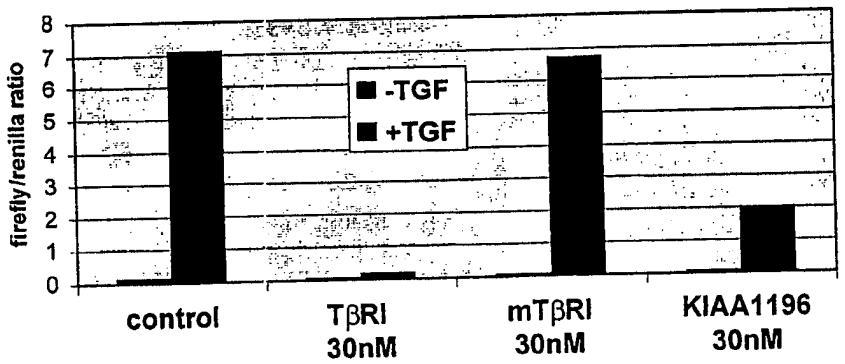


Figure 25: KIAA1196 siRNA specifically represses TGFβ-dependent reporter activity in HEK293 cells

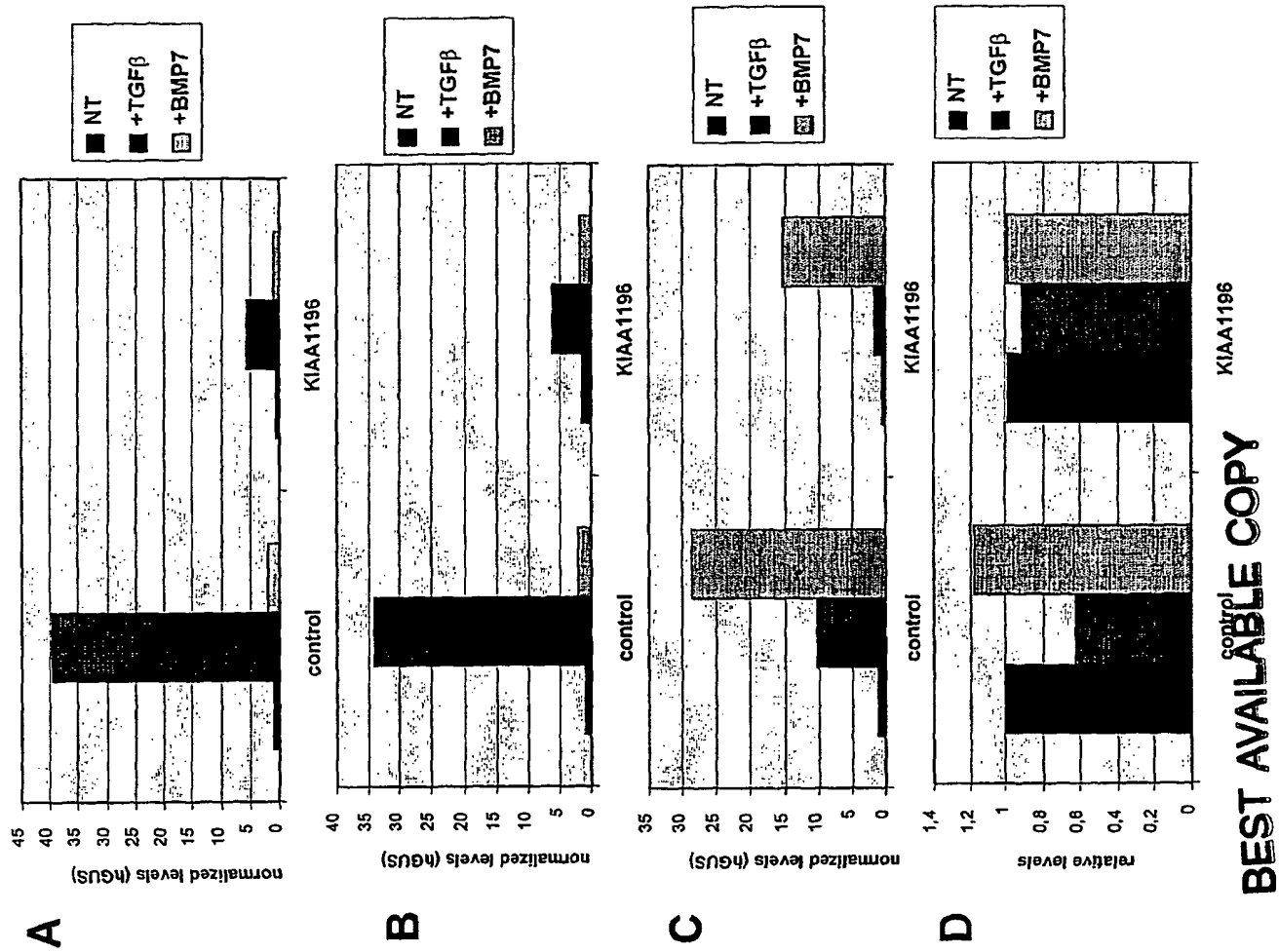
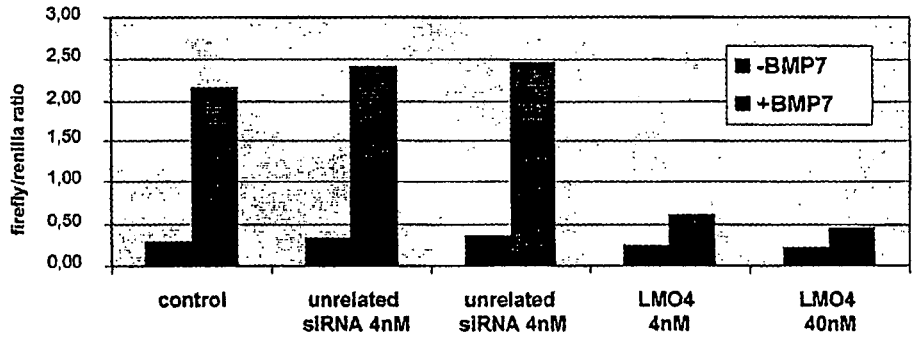


Figure 26: KIAA1196 siRNA specifically represses TGFβ-dependent markers

Figure 27: Interaction between LMO4 and Smad9

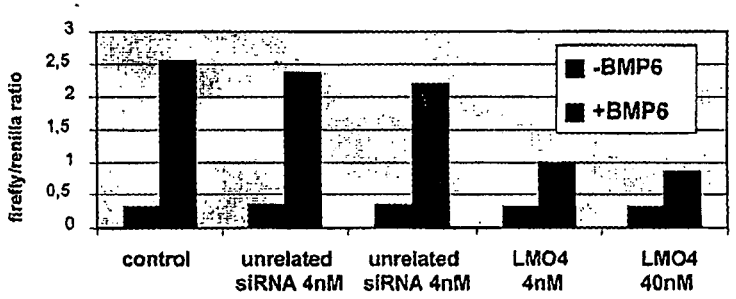
A

LMO4 siRNA regulation of $p(GC)_{12}$-MLP-Luc activity in HepG2 cells



B

LMO4 siRNA regulation of $p(GC)_{12}$-MLP-Luc activity in HepG2 cells



C

LMO4 siRNA regulation of $p(GTCT)_8$-MLP-Luc activity in HepG2 cells

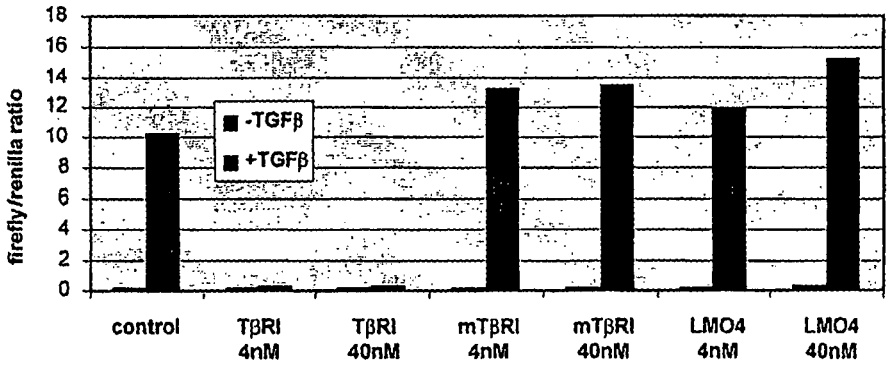
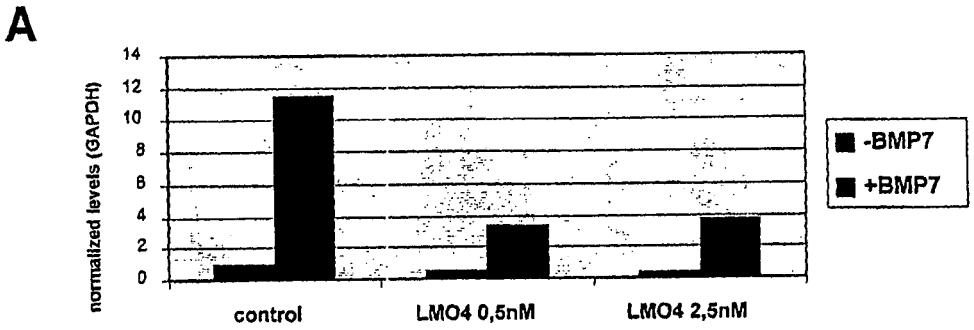


Figure 28:LMO4 siRNA specifically represses BMP-dependent reporter activity

Q-PCR analysis of alkaline phosphatase mRNA levels in HepG2 cells



Q-PCR analysis of alkaline phosphatase mRNA levels in HepG2 cells

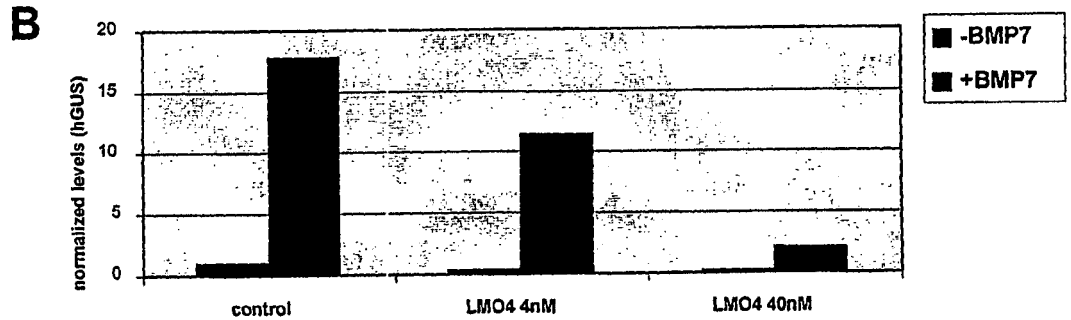
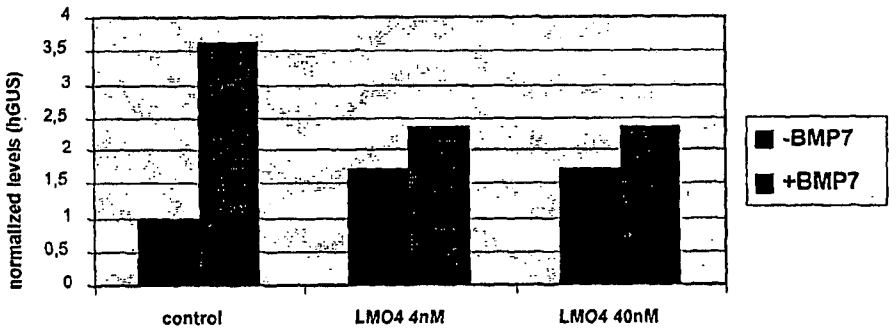


Figure 29:LMO4 siRNA specifically represses BMP-dependent markers

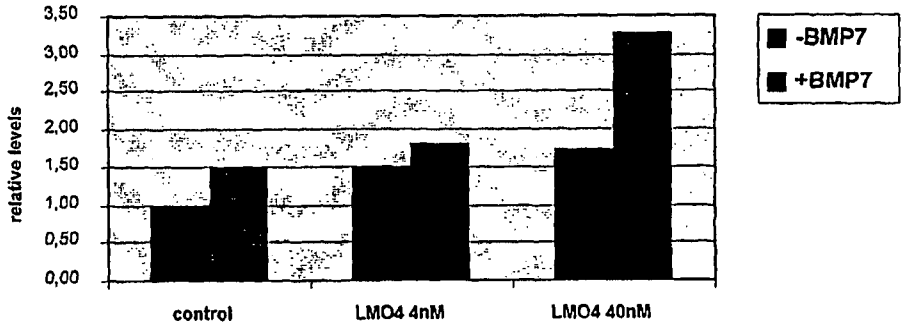
A

Q-PCR analysis of junB mRNA levels in HepG2 cells



B

Q-PCR analysis of hGUS mRNA levels in HepG2 cells



C

Q-PCR analysis of PAI-1 mRNA levels in hepG2 cells

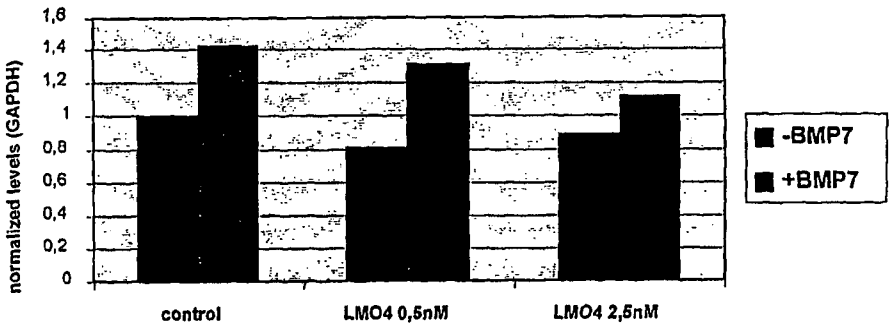


Figure 30:LMO4 siRNA does not repress BMP-independent markers

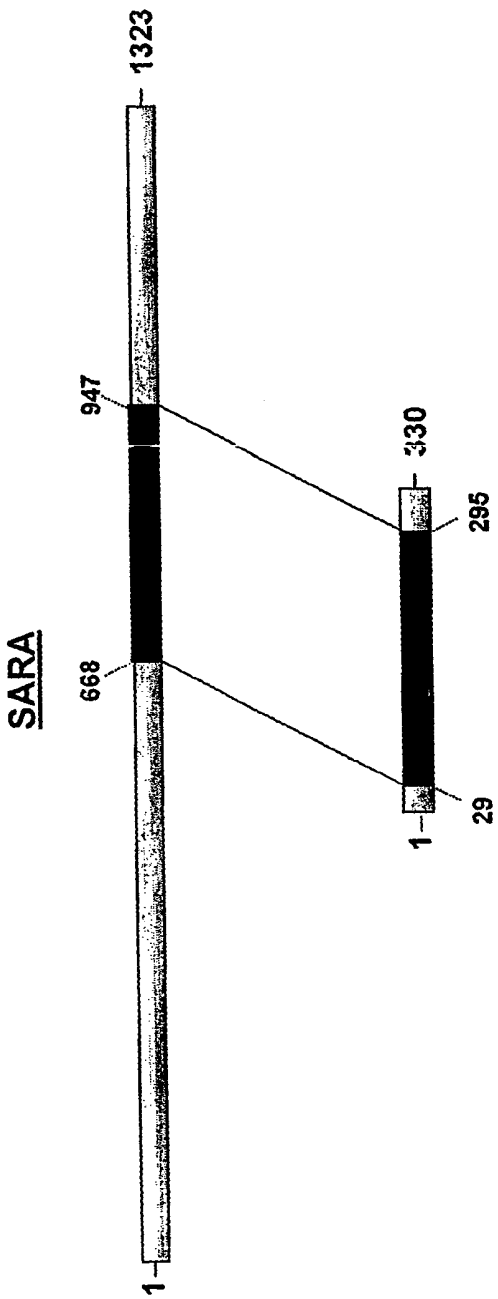
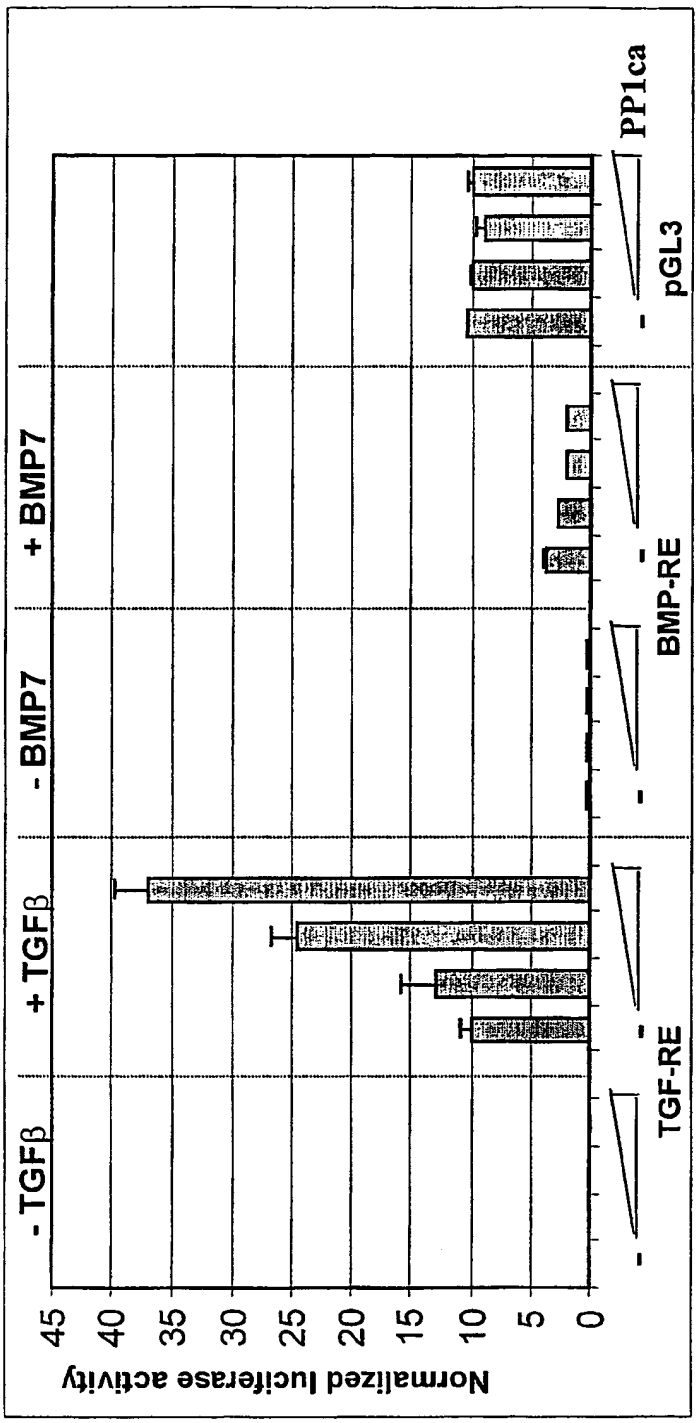


Figure 31: Interaction between PP1ca and SARA



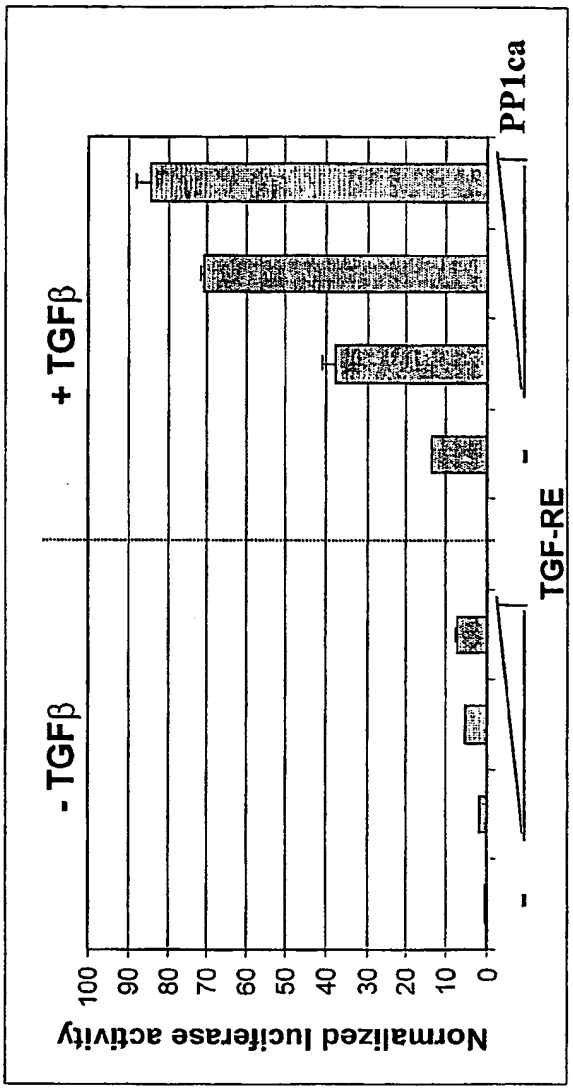


Figure 32: PP1ca stimulates the TGFβ pathway

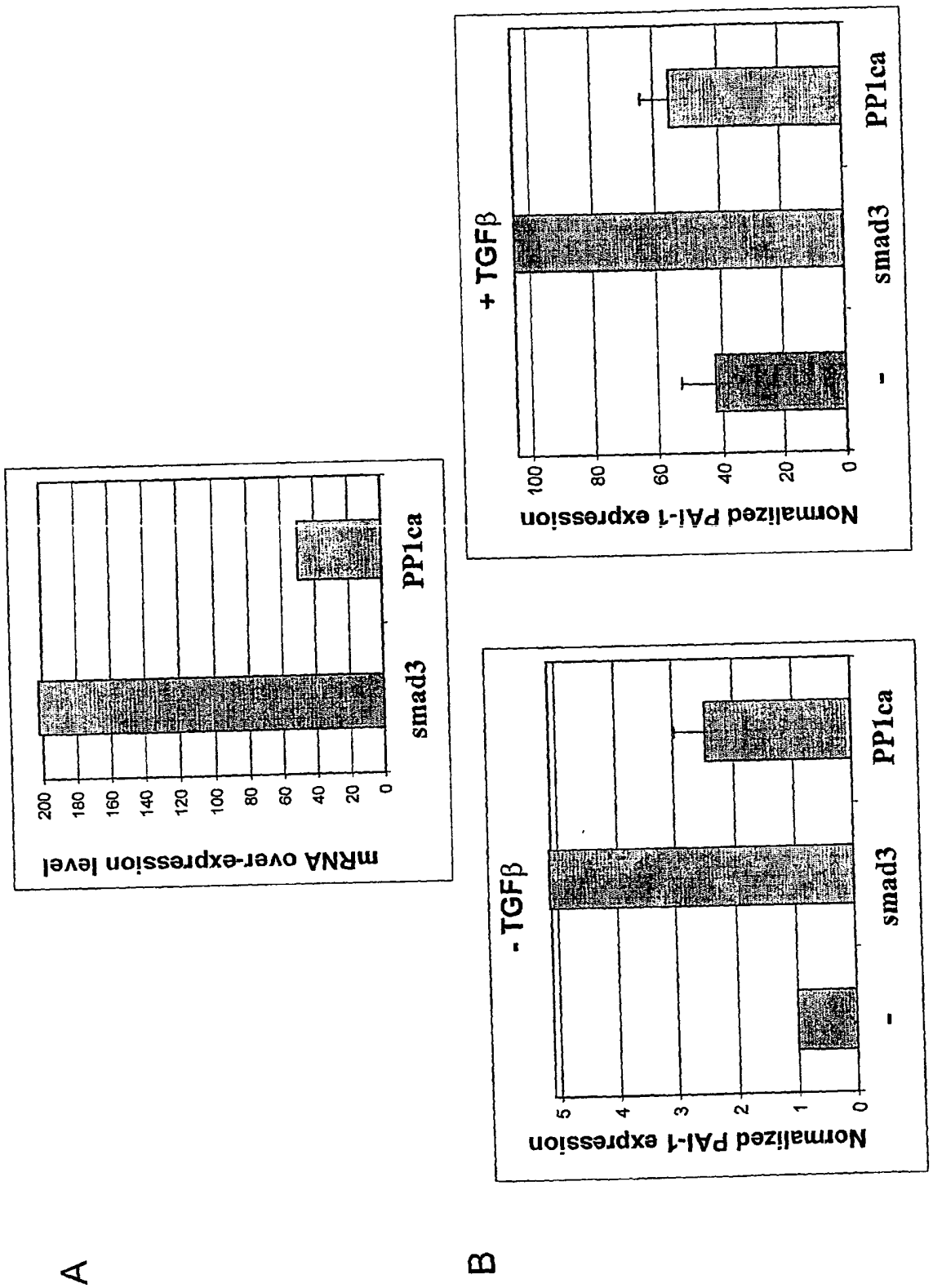


Figure 33: PP1ca stimulates PAI-1 mRNA expression

BEST AVAILABLE COPY

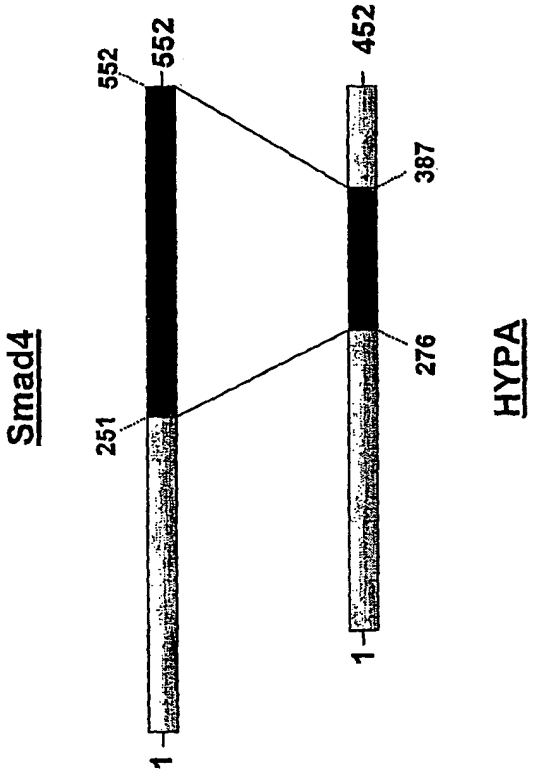


Figure 34: Interaction between HYPA and Smad4

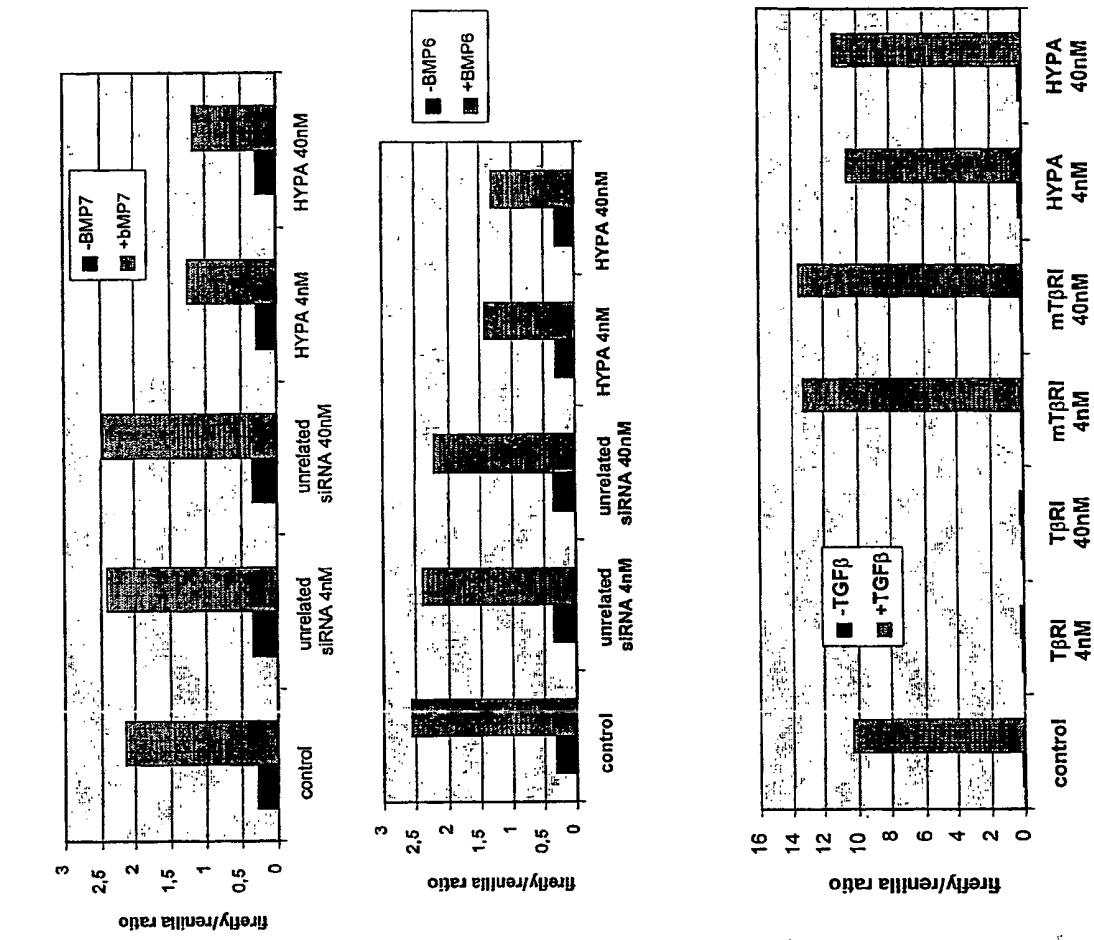


Figure 35: HYPA siRNA specifically represses BMP-dependent reporter activity

BEST AVAILABLE COPY

Q-PCR analysis of alkaline phophatase mRNA levels in HepG2 cells

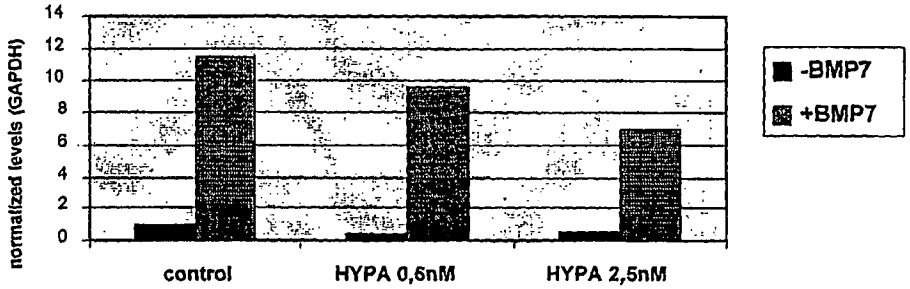


Figure 36: HYPA siRNA represses BMP-dependent markers

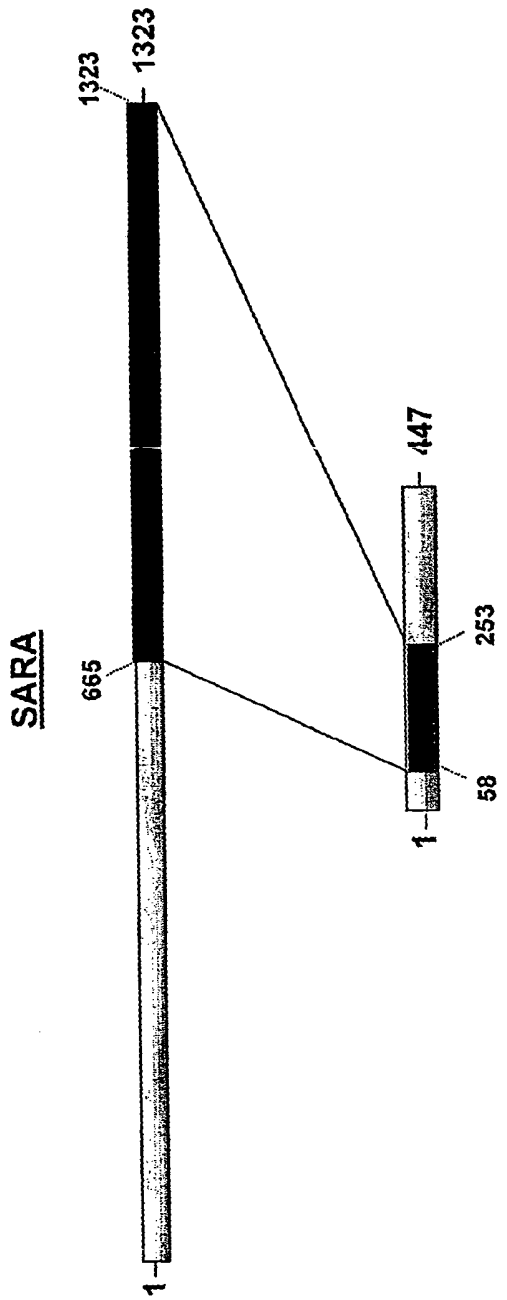


Figure 37: Interaction between FLJ20037 and SARA

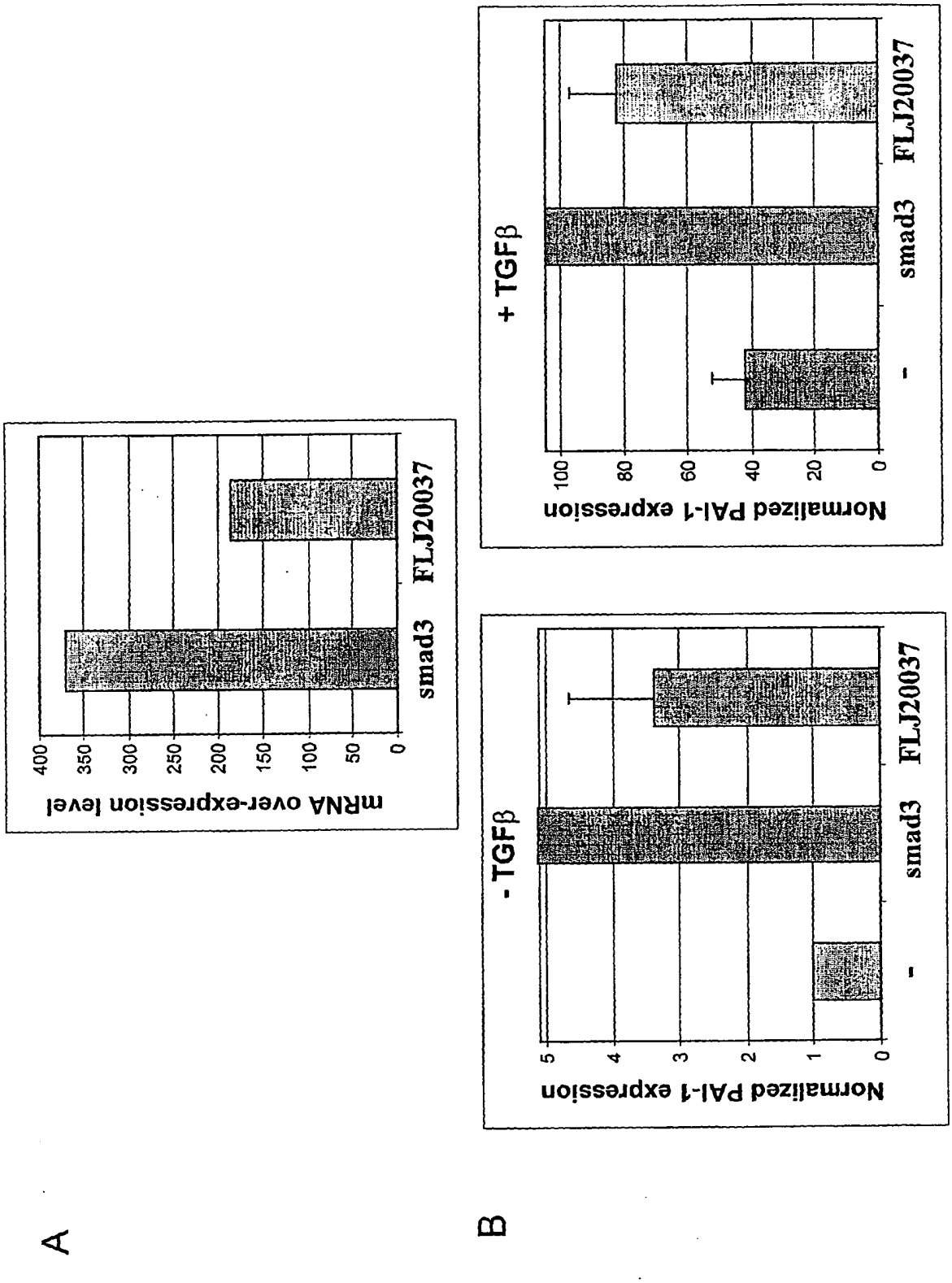


Figure 38: FLJ20037 stimulates PAI-1 mRNA expression

BEST AVAILABLE COPY

Q-PCR analysis of PAI-1 levels in HepG2 cells

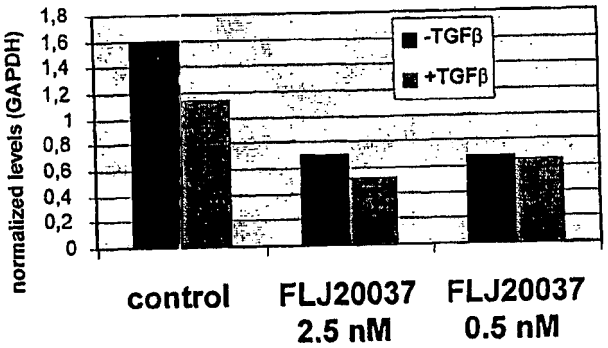


Figure 39: FLJ20037 siRNA down-regulates TGFβ-dependent markers

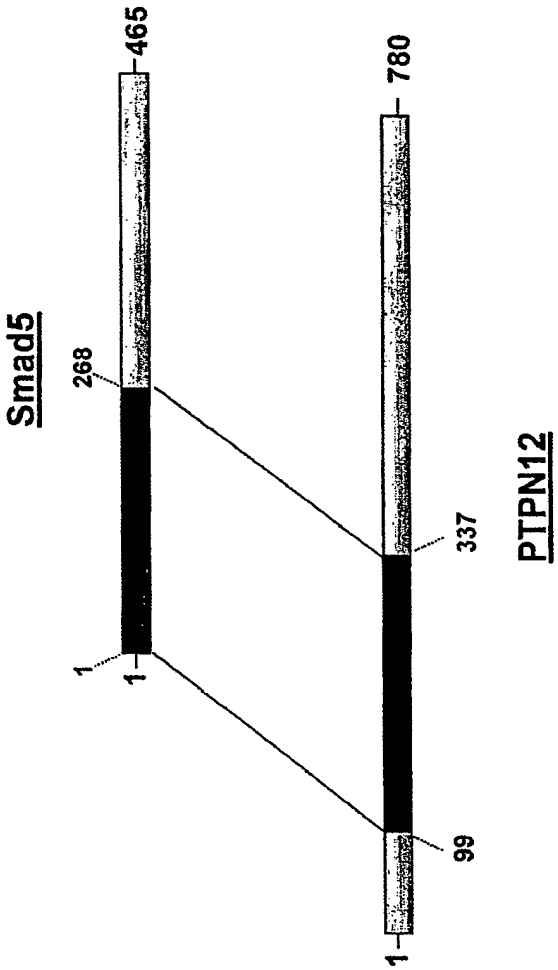
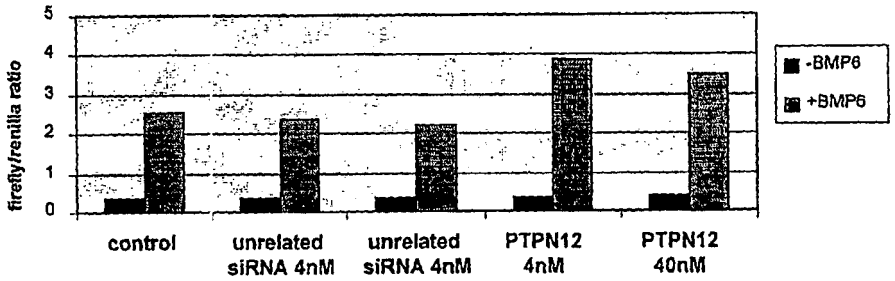


Figure 40: Interaction between PTPN12 and Smad5

A

PTPN12 siRNA regulation of $p(GC)_{12}$-MLP-Luc activity in HepG2 cells



B

PTPN12 siRNA regulation of $p(GTCT)_8$-MLP-Luc activity in HepG2 cells

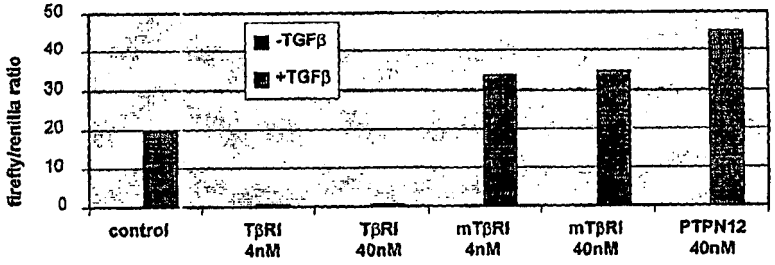
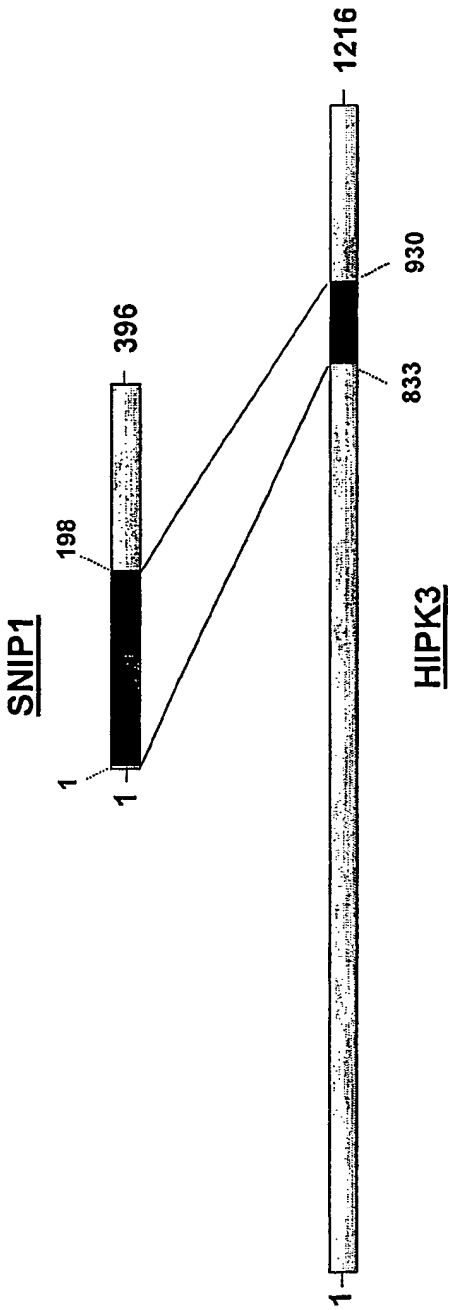
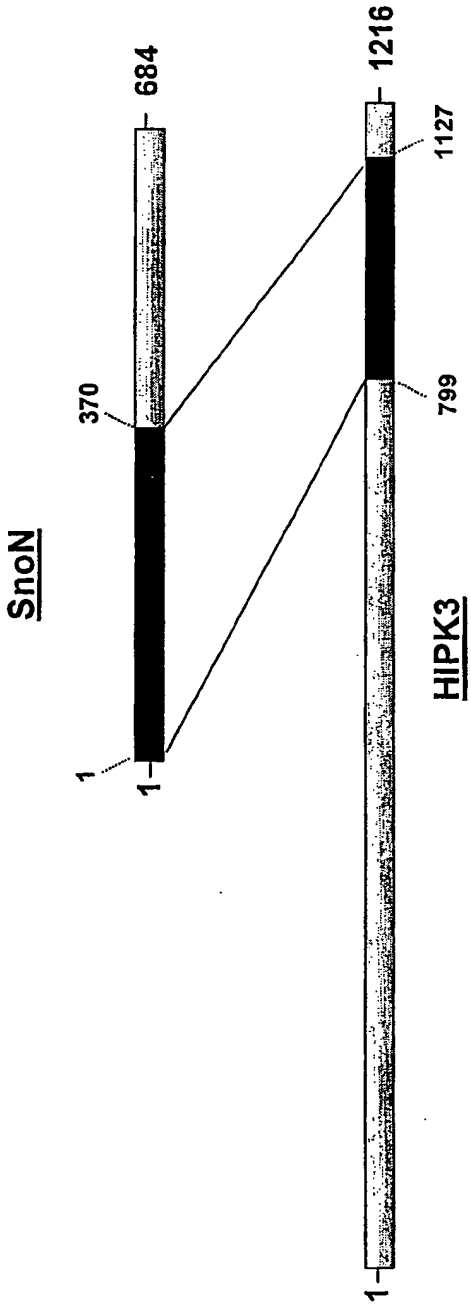


Figure 41: PTPN12 siRNA up-regulates BMP and TGFβ-dependent reporter activities



A

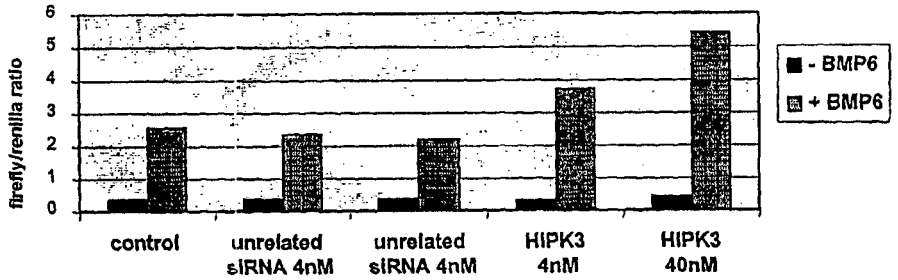


B

Figure 42: Interaction between HIPK3 SnoN and SNIP1

A

HIPK3 siRNA regulation of $p(GC)_{12}$-MLP-Luc activity in HepG2 cells



B

HIPK3 siRNA regulation of $p(GTCT)_8$-MLP-Luc activity in HepG2 cells

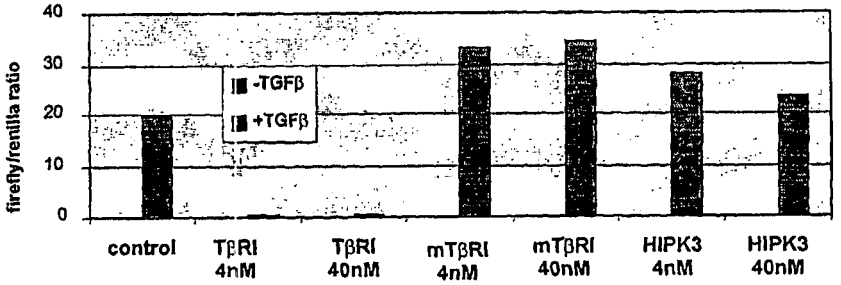


Figure 43: HIPK3 siRNA specifically up-regulates BMP-dependent reporter activities

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
5 June 2003 (05.06.2003)

PCT

(10) International Publication Number
WO 2003/045990 A3

(51) International Patent Classification[7]: C07K 14/46, 14/47, C12N 1/16, 1/21, 5/10, 15/12, 15/63, C12Q 1/68, G01N 33/53, C12N 15/10

(21) International Application Number:
PCT/EP2002/013866

(22) International Filing Date:
26 November 2002 (26.11.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/333,348	26 November 2001 (26.11.2001)	US
60/384,537	31 May 2002 (31.05.2002)	US
60/422,471	30 October 2002 (30.10.2002)	US

(71) Applicant *(for all designated States except US)*: **HYBRIGENICS** [FR/FR]; 3/5 Impasse Reille, F-75014 Paris (FR).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **LEGRAIN, Pierre** [FR/FR]; 5, rue Mizon, F-75015 Paris (FR). **GAUTHIER, Jean-Michel** [FR/FR]; 9 Avenue du Maréchal Galliéni, F-78700 Conflans Sainte Honorine (FR). **COLLAND, Frédéric** [FR/FR]; 16, rue du Manoir, F-95380 Puiseux (FR). **JACQ, Xavier** [FR/FR]; 11 rue de Rambouillet, F-75012 Paris (FR).

(74) Agents: **ERNEST Gutmann-Yves Plasseraud S.A.** et al.; 3 rue Chauveau-Lagarde, F-75008 Paris (FR).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report:
1 April 2004

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: PROTEIN-PROTEIN INTERACTIONS INVOLVING TRANSFORMING GROWTH FACTOR BETA SIGNALLING

(57) Abstract: The present invention relates to protein-protein interactions involved in transforming growth factor β disorders and/or diseases. More specifically, the present invention relates to complexes of polypeptides or polynucleotides encoding the polypeptides, fragments of the polypeptides, antibodies to the complexes, Selected Interacting Domains (SID®) which are identified due to the protein-protein interactions, methods for screening drugs for agents which modulate the interaction of proteins and pharmaceutical compositions that are capable of modulating the protein-protein interactions.

WO 2003/045990 A3

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 02/13866

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C07K14/46 C07K14/47 C12N1/16 C12N1/21 C12N5/10
C12N15/12 C12N15/63 C12Q1/68 G01N33/53 C12N15/10

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, BIOSIS, WPI Data, PAJ, MEDLINE, EMBASE, Sequence Search, SCISEARCH, BIOTECHNOLOGY ABS, CHEM ABS Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 01 64834 A (CHEN RUI HONG ;HYSEQ INC (US); WANG DUNRUI (US); WANG JIAN RUI (US) 7 September 2001 (2001-09-07) page 108, line 12; table 2 page 39, line 27 - page 41, line 5 page 1, line 26 - page 4, line 2	7-15, 17-20
Y	-----	1-4,6,30
Y	WO 98 55512 A (VLAAMS INTERUNIV INST) 10 December 1998 (1998-12-10)	1-4,6,30
A	page 2, line 1 - page 5, line 3 page 10, line 11 - page 14, line 7	7-15, 17-20
	----- -/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
20 August 2003	18. 11. 2003

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	De Kok, A.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 02/13866

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 98 53830 A (HARVARD COLLEGE) 3 December 1998 (1998-12-03) page 2, line 11 - page 12, line 9 page 16, line 17 - page 25, line 2 examples 1-10	1-4, 6-15, 17-20,30

A	LABBE E ET AL: "SMAD2 AND SMAD3 POSITIVELY AND NEGATIVELY REGULATE TGFBETA- DEPENDENT TRANSCRIPTION THROUGH THE FORKHEAD DNA-BINDING PROTEIN FAST2" MOLECULAR CELL, CELL PRESS, CAMBRIDGE, MA, US, vol. 2, no. 1, July 1998 (1998-07), pages 109-120, XP000857311 ISSN: 1097-2765 abstract	1-3

A	MASSAGUE J ET AL: "TGF-BETA SIGNALLING THROUGH THE SMAD PATHWAY" TRENDS IN CELL BIOLOGY, ELSEVIER SCIENCE LTD, XX, May 1997 (1997-05), pages 187-192, XP002911610 ISSN: 0962-8924 the whole document	1

P,X	JIAO KAI ET AL: "Identification of mZnf8, a mouse Kruppel-like transcriptional repressor, as a novel nuclear interaction partner of Smad1." MOLECULAR AND CELLULAR BIOLOGY, vol. 22, no. 21, November 2002 (2002-11), pages 7633-7644, XP002251797 November, 2002 ISSN: 0270-7306 cited in the application the whole document	1-3,20, 30

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP 02/13866

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 5 and 16 completely and 1-4, 6 and 30 partially
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-4,6-15,17-19,30 all partially and 20 completely

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 5 and 16 completely and 1-4, 6 and 30 partially

Present claims 1-3 and 30 relate to all possible complexes between to lists of proteins idetified in column 1 resp. column 4 of Table 2. Support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT is to be found, however, for only a very small proportion of the complexes claimed. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Consequently, the search has been carried out for those parts of the claims which appear to be supported and disclosed, namely those parts relating to the complexes specifically disclosed in Table 2.

Present claims 4 and 6 relate to the use of a "SID", an "interaction" or a "prey" to screen for inhibitors of the TGFb signalling pathway, without giving any description in technical terms of such "SID", "interaction" or "prey". Therefore, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Consequently, the search has been carried out for those parts of the claims which appear to be supported and disclosed, namely those parts relating to the SID's defined in claim 7 resp. claim 8.

Present claims 5 and 16 relate to a compound defined by reference to a desirable characteristic or property, namely by being an inhibitor of the TGFbeta signalling pathway.

The claims cover all compounds having this characteristic or property, whereas the application provides support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for NONE of such compounds. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the compound by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, NO search has been carried out for those claims.

The applicant's attention is drawn to the fact that claims relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

claims are amended following receipt of the search report or during any Chapter II procedure. If the application proceeds into the regional phase before the EPO, the applicant is reminded that a search may be carried out during examination before the EPO (see EPO Guideline C-VI, 8.5), should the problems which led to the Article 17(2) declaration be overcome.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-4, 6-15, 17-19, 30 all partially and 20 completely

A complex between repectively human proteins smad1 or smad4 or smad5 or smad9a and human protein ZNF8; a complex between human nucleic acids encoding proteins smad1 or smad4 or smad5 or smad9a and human nucleic acid encoding protein ZNF8; a specific interacting domain (SID) derived from said ZNF8 protein; variants of said SID's; vectors comprising polynucleotides encoding said SID's; pharmaceutical compositions comprising said SID's; a protein chip comprising the proteins identified above and use of said ZNF8 or said complexes for the treatment of TGFbeta related disorders.

---

2. claims: 1-4, 6-15, 17-19, 30 all partially and 21 completely

A complex between human protein smurf2 and human protein LAPTm5; a complex between a human nucleic acid encoding the protein smurf2 and human nucleic acid encoding the protein LAPTm5; a specific interacting domain (SID) derived from said LAPTm5 protein; variants of said SID's; vectors comprising polynucleotides encoding said SID's; pharmaceutical compositions comprising said SID's; a protein chip comprising the proteins identified above and use of said LAPTm5 or said complexes for the treatment of TGFbeta related disorders.

---

3. claims: 1-4, 6-15, 17-19, 30 all partially and 22 completely

A complex between repectively human proteins smurf2 or sara and human protein RNF11; a complex between human nucleic acids encoding proteins smurf2 or sara and human nucleic acid encoding protein RNF11; a specific interacting domain (SID) derived from said RNF11 protein; variants of said SID's; vectors comprising polynucleotides encoding said SID's; pharmaceutical compositions comprising said SID's; a protein chip comprising the proteins identified above and use of said RNF11 or said complexes for the treatment of TGFbeta related disorders.

---

4. claims: 1-4, 6-15, 17-19, 30 all partially and 23 completely

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

A complex between human protein smad9 and human protein LM04; a complex between a human nucleic acid encoding the protein smad9 and human nucleic acid encoding the protein LM04; a specific interacting domain (SID) derived from said LM04 protein; variants of said SID's; vectors comprising polynucleotides encoding said SID's; pharmaceutical compositions comprising said SID's; a protein chip comprising the proteins identified above and use of said LM04 or said complexes for the treatment of prostate cancer.

---

5. claims: 1-4, 6-15, 17-19, 30 all partially and 24 completely

A complex between human protein sara and human protein PPC1; a complex between a human nucleic acid encoding the protein sara and human nucleic acid encoding the protein PPC1; a specific interacting domain (SID) derived from said PPC1 protein; variants of said SID's; vectors comprising polynucleotides encoding said SID's; pharmaceutical compositions comprising said SID's; a protein chip comprising the proteins identified above and use of said PPC1 or said complexes for the treatment of TGFbeta related disorders.

---

6. claims: 1-4, 6-15, 17-19, 30 all partially and 25 completely

A complex between human protein smad4 and human protein HYPA; a complex between a human nucleic acid encoding the protein smad4 and human nucleic acid encoding the protein HYPA; a specific interacting domain (SID) derived from said HYPA protein; variants of said SID's; vectors comprising polynucleotides encoding said SID's; pharmaceutical compositions comprising said SID's; a protein chip comprising the proteins identified above and use of said HYPA or said complexes for the treatment of TGFbeta related disorders.

---

7. claims: 1-4, 6-15, 17-19, 30 all partially and 26 completely

A complex between human protein smad5 and human protein PTP; a complex between a human nucleic acid encoding the protein smad5 and human nucleic acid encoding the protein PTP; a specific interacting domain (SID) derived from said PTP protein; variants of said SID's; vectors comprising polynucleotides encoding said SID's; pharmaceutical compositions comprising said SID's; a protein chip comprising the proteins identified above and use of said PTP or said complexes for the treatment of TGFbeta related disorders.

---

8. claims: 1-4, 6-15, 17-19, 30 all partially and 27 completely

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

A complex between repectively human proteins snipl or snon and human protein HYPK3; a complex between human nucleic acids encoding proteins snipl or snon and human nucleic acid encoding protein HYPK3; a specific interacting domain (SID) derived from said HYPK3 protein; variants of said SID's; vectors comprising polynucleotides encoding said SID's; pharmaceutical compositions comprising said SID's; a protein chip comprising the proteins identified above and use of said HYPK3 or said complexes for the treatment of TGFbeta related disorders.

---

9. claims: 1-4, 6-15, 17-19, 30 all partially and 28 completely

A complex between human protein smad1 and human protein KIAA1196; a complex between a human nucleic acid encoding the protein smad1 and human nucleic acid encoding the protein KIAA1196; a specific interacting domain (SID) derived from said KIAA1196 protein; variants of said SID's; vectors comprising polynucleotides encoding said SID's; pharmaceutical compositions comprising said SID's; a protein chip comprising the proteins identified above and use of said KIAA1196 or said complexes for the treatment of TGFbeta related disorders.

---

10. claims: 1-4, 6-15, 17-19, 30 all partially and 29 completely

A complex between human protein sara and human protein FL20037; a complex between a human nucleic acid encoding the protein sara and human nucleic acid encoding the protein FL20037; a specific interacting domain (SID) derived from said FL20037 protein; variants of said SID's; vectors comprising polynucleotides encoding said SID's; pharmaceutical compositions comprising said SID's; a protein chip comprising the proteins identified above and use of said FL20037 or said complexes for the treatment of TGFbeta related disorders.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/EP 02/13866

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0164834 A	07-09-2001	AU 4151101 A	12-09-2001
		CA 2401505 A1	07-09-2001
		EP 1276751 A2	22-01-2003
		WO 0164834 A2	07-09-2001
		AU 3834701 A	12-09-2001
		AU 3995501 A	12-09-2001
		AU 4187301 A	12-09-2001
		WO 0164835 A2	07-09-2001
		WO 0164839 A2	07-09-2001
		WO 0164840 A2	07-09-2001
WO 9855512 A	10-12-1998	AU 8623798 A	21-12-1998
		WO 9855512 A2	10-12-1998
		EP 0994945 A2	26-04-2000
		JP 2002505577 T	19-02-2002
		US 6313280 B1	06-11-2001
		US 2002035246 A1	21-03-2002
WO 9853830 A	03-12-1998	AU 7705398 A	30-12-1998
		US 2002160355 A1	31-10-2002
		US 6365711 B1	02-04-2002
		WO 9853830 A1	03-12-1998